Information Aggregation in Directed Networks*

Christopher Brown† Svetlana Pevnitskaya‡

August 29, 2024

Abstract

Individuals often can obtain more information by observing previous decisions of others. We present a framework of individual choice under uncertainty for agents connected via a directed network which allows for observational learning. The decision is made once, thus learning from repetition is not possible. We outline properties of networks that affect accuracy of individual choice and information aggregation. Performance is evaluated using two criteria: individual (final agent) and social (group) choice accuracy, with the result that network properties that enhance performance under one criterion reduce performance under the other. We test theory by designing an experiment with two treatment variables: (1) network structure and (2) method of assigning subjects to positions within a network. In all treatments, there is efficiency loss compared to a benchmark with all Bayesian agents. On average individuals understand the value of observational learning, however there is heterogeneity in the willingness to pay for later positions. In networks with endogenous assignment, the advantage of later positions is eroded as a result of self-selection of more rational subjects to those positions, and greater noise in choices of subjects in earlier positions. Persons exhibit a propensity to overweight information inferred from observed actions of others as opposed to own private signal, which is increasing in the number of observed actions.

Keywords: Networks, Social Learning, Information Cascades, Experiment

JEL classification codes: C91, C92, D83, D85

*We thank participants of seminars at NYU, Caltech, CGU, North American Summer Meeting of Econometric Society (NASMES), Economic Science Association (ESA), Western Economic Association International (WEAI) and Bay Area Behavioral and Experimental Economics Workshop (BABEEW) meetings, Andrew Caplin, Guillaume Frechette, Dorothea Kübler, Georg Weizsäcker, Marco Battaglini, Matthew Jackson and others for helpful comments.

†E-mail: `cbrown.econ@gmail.com`

‡E-mail: `spevnitskaya@gmail.com`

1 Introduction

In many social and economic settings decisions are made under imperfect information about the payoffs of available actions. Choice under uncertainty problems have been studied extensively in economics and related fields, with the standard analysis focusing on individual decision making. This approach does not take into account the fact that choices are rarely made in isolation. In many settings, prior to taking an action, a person is able to infer additional information by observing actions previously taken by others. Individual actions then, in addition to generating private payoffs, affect social learning. This act of learning through observation may allow privately held information to be aggregated, in which case individuals acting later make more informed choices. The decision problem faced by each agent becomes different depending on the information they have at the time of making the decision, and requires extending individual decision theory to account for observational social learning.

There are numerous examples of settings where individuals take into account choices of others. As one example, consider a consumer who is deciding between the latest models of competing brands of cell phones. If he is uncertain about which alternative is better, he may note purchasing decisions made by friends and family. Choices of others, particularly if most of them select the same brand, may change the persons' original prior about the products and lead to a decision to follow suit. As another example, consider a person who has to choose among alternative surgical procedures, a single irreversible choice, with the best match and thus the success of the treatment becoming clear only after the procedure. A person may want to explore choices of others taken under similar circumstances albeit without fully knowing their private medical information. This setting corresponds to our framework of directed networks since others have already made their decisions. Our results on the impact of network structure on choice accuracy inform how the person should collect information about previous choices of others, and can also be used to assess whether information aggregation within the structure of the organization is optimal for the CEO or the Chair who is making the final and most significant decision.

The nature of information transmission throughout a group, organization, or society plays a key role in this interdependence of decisions. Prior research (see, e.g., Golub and Jackson (2010)) has identified social networks as the primary channel through which information and opinions are exchanged. The structure of the network in which the individuals are embedded determines whose actions can be observed and therefore, what information can be obtained via observation. Social, economic, and organizational networks are complex and have various structures which lead to different opportunities for social learning. In this paper we set out to identify properties of networks that make observational learning successful and increase the likelihood of selecting an objectively better option. We evaluate the impact such properties on improving the choices of the final agent in the network and also the overall accuracy of the whole group.

Our theoretical framework includes a set of Bayes-rational agents, who are uncertain about the payoffs of their possible choices, and sequentially make a single, irreversible decision. Each agent has a set of available actions, each corresponding to a possible state, and earns positive payoff only if their chosen action matches the true state, which is unknown at the time of making a decision and is revealed afterwards. Each individual receives an imperfectly informative private signal, and, prior to acting, observes a subset of actions taken by other agents. That observed subset depends on the structure of the network and the individual's position within it. Since each agent is paid based on own action, it reveals their beliefs, and possibly their private independent signal.[1]

We introduce the concept of the informativeness of actions within the network, define its degrees, and obtain correspondence between the structure of the network and the informativeness of individual actions within it. The optimal decision rule takes into account whether observed actions inform of private signals, which is a function of the network and the specific composition of the set of observed prior actions. When enough prior actions have been observed, an individual may fall into an information cascade and rationally disregard own private information in favor of imitating the action taken by the majority of those they observe. In this case, the individual's action no longer conveys information about their private signal. Thus, the formation of a cascade limits the aggregation of privately held information and reduces the externality benefits which would otherwise accrue to subsequent decision makers. We present an individual decision rule which establishes precisely how individuals' optimal choice depends on own private signal and information inferred through observed actions.

Network performance is evaluated using two criteria: an *individual criterion* which considers the expected payoff of the final agent and a *social criterion* which takes a standard utilitarian approach in comparing the aggregate payoff of all agents in the network. We show that there is a tradeoff between network properties which improve decision accuracy under each criteria. Properties that produce the highest accuracy for the group overall, reduce the payoff to the final agent, and vice versa.

[1]Unlike individual decisions under uncertainty, in this environment an action taken by the agent is observed by other agents and affects their decision making process. One can think of this interdependence as an information externality. However, unlike, for example, environments characterized by consumption externalities (See, e.g., Katz and Shapiro (1985)) in which individual's utility from a particular action depends, at least in part, upon how many others are choosing the same action), here the interdependence is different: an agent's payoff depends only on the accuracy of own action, regardless of others' choices. In our setting an agent experiences positive externality, only if observed action reveals signal.

In order to test our theoretical results, we design an experiment with two main treatment variables: the structure of the network and the assignment to positions within the network. For the first treatment variable variation we use network structures that benefit the performance of the final agent vs of the whole group. As a second treatment variation, we employ endogenuos vs exogenous assignment to positions. In organizational structures an individual may be interested in taking a specific position within it, and furthermore there might be heterogeneity in the assessments of the benefits of the positions. We therefore want to investigate if there is self-selection to positions, if it is based on heterogeneity in willingness to pay for positions, and whether it would affect observational learning. The second treatment variable allows us to explore further any behavioral deviations from theoretical predictions and to provide insights on the observational learning in the field. To highlight observational learning as a focus of this study, and give it the best shot, we simplify the nature of uncertainty and consider a setting with binary states and actions.

We find that there are moderate efficiency losses compared to a benchmark in which all individuals are Bayes-rational. Overall, about 85% of choices are consistent with the benchmark of full Bayesian rationality. The efficiency loss is driven by behavioral phenomena which largely amounts to not weighting own independent private information enough and imitating observed choices more than it is optimal. Compared to Bayesian benchmark, subjects decrease weight assigned to own private signal as the number of observed actions increases. Specifically, in cases of equal posterior beliefs, subjects tend to go with own signal when the number of observed actions is low and against own signal as the number of observed actions increases. We estimate subjects' demand for information in treatments with endogenous assignment to network positions by designing an auction mechanism that reveals their valuations of positions' premiums. Our data indicate that individuals understand the value of information in this environment – subjects tend to bid higher for positions with more observed actions. However, we also find that in endogenous assignment treatments, the frequency of choices consistent with Bayesian is skewed towards later positions, indicating that more rational subjects who understand the benefits of larger observation sets self-select themselves to such positions. This leads to more "errors" in the earlier positions and erodes the payoff to the final agent. Overall experimental data supports theoretical results on the effect of network structure on the performance of the whole group.

The remainder of the paper is organized as follows. Section 2 provides an overview of the relevant theoretical and experimental literature and describes how the present study complements and adds to the existing body of knowledge. The model and theoretical results are presented in Section 3. We describe experimental design, selection of parameters, and the procedures of our laboratory experiment in Section 4. Our results are discussed in Section

5. Section 6 concludes.

2 Related Literature

2.1 Theoretical

The study of observational learning began with the seminal works of Banerjee (1992) and Bikhchandani et al. (1992) which independently sought to explain observed uniformity in social behavior.[2] These early studies assumed that agents make decisions in a linear sequence and each agent is able to observe the full history of actions. Since, in most cases, individuals have much less complete information, some follow up studies, which were still based on the sequential process, assumed, for example, that a random sample of past actions is observed (Banerjee and Fudenberg, 2004), or that individuals only observe the action of their immediate predecessor (Çelen and Kariv, 2004b). A recent strand of literature extends the framework by considering models in which agents are embedded in a network of connections and observe a subset of prior actions that is determined by the structure of this network. Golub and Sadler (2016) refer to models of this type, collectively, as the *sequential social learning model* (hereafter, SSLM).

Acemoglu et al. (2011), Lobel and Sadler (2015), and Arieli and Mueller-Frank (2018) all consider a SSLM on a stochastically generated network and examine how properties of these networks affect information aggregation and social learning. Yet, these analyses differ from the current study in that their focus is on studying asymptotic learning in very large or infinite network structures.[3] While asymptotic results are crucial for understanding the dynamics of learning in larger societies, they may fail to describe some intricacies of the choice environments of smaller social and organizational groups. Our analysis, on the other hand, examines social learning and information aggregation within relatively small groups where the effects of network structure are less understood[4] while an individual might have greater control or ability to change the structure of smaller groups.

Out study is also different from settings in which individuals form beliefs and make decisions over multiple periods (Gale and Kariv, 2003; Golub and Jackson, 2010, 2012; Mossel

[2] A subsequent, more general analysis is presented in Smith and Sørensen (2000). More recently, Rosenberg and Vieille (2019) revisit this environment in order to study the efficiency of social learning under various conditions.

[3] Asymptotic learning is said to have occurred if the proportion of agents choosing the optimal action converges to one over time.

[4] Golub and Sadler (2016) note that "Asymptotic outcomes of sequential observational learning are now well understood, but important challenges remain. A significant gap in our knowledge concerns short-run dynamics and rates of learning in these models."

et al., 2015). While these studies are complementary in their investigation of the ability to update beliefs, they assume that choices are made repeatedly and that the decision maker has the opportunity to revise actions over time, which does not accurately reflect the decision environment of many important choices.[5]

2.2 Experimental

The earliest analysis of observational learning in a laboratory setting is due to Anderson and Holt (1997). Following early theory literature, their experimental analysis considers the baseline observational learning environment in which individuals observe the entire history of actions. Their results indicate that subjects typically use their available information in a rational, Bayesian manner and that information cascades occur regularly. Several subsequent experimental studies have raised doubt about individuals' ability to properly reason in a Bayesian manner. For example, experimental subjects often display limited depth of reasoning (Kübler and Weizsäcker, 2004) and have a tendency to overvalue their private information relative to information inferred from observed actions (Nöth and Weber, 2003; Weizsäcker, 2010). As a result of these behavioral aspects of individual choice, cascade formation may often be delayed beyond its theoretically predicted starting point and cascades may be short-lived or even reversible (Goeree et al., 2007).

Experimental studies examining the role of network structure within a social learning environment have thus far been limited to environments with repeated learning over multiple time periods. This is a different setting than the one considered here. In these studies agents make decisions simultaneously, then the choices of neighbors are revealed, followed by repetition of the choice, etc., in order to analyze the convergence of beliefs. Choi et al. (2005, 2012) report the results of an experiment designed to test the predictions of the model of Gale and Kariv (2003) in simple three player networks. Grimm and Mengel (2020) conduct a similar experiment but employ several, more complex seven-player networks. Chandrasekhar et al. (2019) report the results of two experiments designed to test different models of social learning in networks in this repeated learning environment. To the best of our knowledge, ours is the first study to experimentally examine the SSLM under one-shot sequential learning with varying network conditions.

[5] Thaler (2016) notes that "Few of us buy cars often enough to get very good at it, and the really big decisions like careers, marriages, and retirement saving give very little room for learning...Either the real world is mostly high stakes or it offers myriad opportunities to learn–not both." Indeed, there are many decisions that are irreversible and cannot be taken again in identical settings, for example having a surgery or going to war.

3 Theoretical Framework

3.1 Choice Environment and Network Structures

There is a finite set of agents $N = \{1, 2, \ldots, n\}$, $n \geq 2$, connected in a network structure, G. Let $\Omega = \{A, B\}$ be the set of possible, payoff relevant states and $\omega \in \Omega$ denote the unknown, underlying state of the world. We index agents according to the order in which they make decisions. Each agent i makes a single, irreversible decision, denoted by $x_i \in \{A, B\}$, which is their prediction regarding the true state. Individuals receive a fixed payoff $V > 0$ if their choice corresponds to the true state and nothing otherwise. Thus, each agent's preferences can be represented by

$$u_i(x_i, \omega; G) = \begin{cases} V, & \text{if } x_i = \omega \\ 0, & \text{otherwise} \end{cases}$$

For simplicity, we will assume throughout that each state is, a priori, equally likely and this is common knowledge (i.e., there is a common prior $p_0 = P(\omega = A) = P(\omega = B) = 1/2$).[6] Each agent receives a private signal, $s_i = \{a, b\}$, that is imperfectly informative with precision $q = P(a \mid \omega = A) = P(b \mid \omega = B) \in (1/2, 1)$. Signals are independent conditional upon the true state. Additionally, agents observe the actions (but not signals) of others that act before them and who they are connected to via the network structure.

Networks are represented by a collection of links which are listed as subsets of N of size 2. We write $ij \in G$ to indicate that the link from i to j exists in network G. Given, the sequential nature of the choice environment, networks are assumed to be *directed* with each agent k observing the action of every agent i for which there exists a *directed path* from i to k in the network.[7] Furthermore, we restrict attention throughout to a subset of directed networks which we refer to as *sequential networks*.

Definition 1 (Sequential network) *A directed network G is a sequential network if*

i) for all $i \in N$, there is no $j < i$ such that $ij \in G$

ii) for all $i \in \{1, 2, \ldots, n-1\}$ there is a unique $j > i$ such that $ij \in G$

This definition implies that the action of every non-final agent is directly observed by only one other agent. Therefore, the decision making process eventually concludes with a unique

[6] This is a prevalent assumption in the existing literature on observational learning–both theoretical (Bikhchandani et al., 1992; Smith and Sørensen, 2000; Acemoglu et al., 2011) and experimental (Anderson and Holt, 1997; Çelen and Kariv, 2004b).

[7] We consider a class of *directed* networks such that if $ij \in G$ then $ji \notin G$. If we denote a link ij by $\ell_i \ell_j$, then a *directed path* in network G from i to k is a set of links $\ell_1\ell_2, \ell_2\ell_3, \ldots, \ell_{M-1}\ell_M$ such that $\ell_m \ell_{m+1} \in G$ for each $m \in \{1, 2, \ldots, M-1\}$ with $\ell_1 = i$ and $\ell_M = k$, and where each node $\ell_1, \ell_2, \ldots, \ell_M$ is distinct.

final agent. This allows us to apply findings to organizations with a leading decision maker such as a CEO, Chairperson, commander-in-chief, general, etc. This definition, additionally, eliminates the possibility of networks in which any agent or set of agents is isolated (i.e., we consider directed networks comprised of a single, weakly connected component). Since our focus is on examining how the structure of a network affects the aggregation of information and observational learning, and unconnected agents are described by standard individual decision theory, we ignore such cases in this paper.

In sequential networks, the set of actions observed by each agent i, which we denote by $O_i \subseteq \{x_1, x_2, \ldots, x_{i-1}\}$, is entirely determined by the structure of the network G. Since individual decisions take into account information that may be inferred from observed actions, it will be useful to refer to the information structure generated by a specific network structure.

Definition 2 (Information Structure) *The information structure of a sequential network G is $I(G) \equiv \{O_1, O_2, \ldots, O_n\}$*

Since much of the observational learning literature has focused on a simple linear network in which each agent observes the full history of actions taken prior to her when making a decision, we select this network to serve as a benchmark and refer to it as the *complete history network*.

Definition 3 (Complete History Network) *The complete history network, denoted by G^C, has $ij \in G^C$ if and only if $j = i + 1$*

In the complete history network, each agent is directly connected to their immediate successor in the sequence of decision makers, and, for any agents i and j with $i > j$, there exists a directed path from j to i. Therefore, each agent observes the actions of every agent who acts before them in the sequence, and G^C generates an information structure $I(G^C)$ with $O_1 = \varnothing$ and $O_i = \{x_1, x_2, \ldots, x_{i-1}\}$ for all $i \in \{2, \ldots, n\}$.

3.2 Individual Choice and Information Cascades

When making a decision, each agent i observes own independent private signal, s_i, and O_i. We denote the information set of agent i by $I_i = \{s_i, O_i\}$. Our benchmark solution concept is that each agent uses their available information and Bayes' rule to form a posterior belief, $p_i = P(\omega = A \mid I_i)$. Since agents' preferences are such that they want their action to correspond to the true state, a rational Bayesian agent will choose $x_i = A$ if $p_i > 1/2$ and $x_i = B$ if $p_i < 1/2$. If $p_i = 1/2$, i.e. an individual is indifferent between alternatives, we assume a tie-breaking rule that leads them to side with their private independent information, s_i.[8]

Assumption 1 *An agent i with posterior belief $p_i = P(\omega = A \mid I_i) = 1/2$, chooses $x_i = A$ if $s_i = a$ and $x_i = B$ if $s_i = b$*

It is readily apparent that any individual with $O_i = \varnothing$ (i.e., one that observes no prior actions) will choose the action that corresponds to private signal. Furthermore, any agent with $|O_i| = 1$ observes only a single prior action and, as a consequence of the tie-breaking rule, will also choose the action corresponding to own private signal regardless of the observed action. The choice for individuals with $|O_i| \geq 2$ is not determined ex ante as it can be influenced by the actions they observe.

When enough prior actions have been observed (two or more in the case of binary states), it is possible for an information cascade to form. This occurs when, given the information inferred from observed actions, it is optimal for an individual to disregard own private signal and imitate the majority action taken by those he observes. Information cascades therefore can be started with very little information and may result in herding on either the correct or incorrect state. Formally, we say that an individual is in a cascade if, for any $s_i \in \{a, b\}$, $P(\omega = z \mid O_i) > \frac{1}{2}$ and $P(\omega = z \mid I_i) > \frac{1}{2}$ hold for some $z \in \{A, B\}$.

Prior to the formation of a cascade, individuals can perfectly infer the private signal of another individual by observing their choice. Such actions are said to be *informative*. However, if an individual is in a cascade, their action no longer conveys their signal. Given this insight, we follow the approach of Bikhchandani et al. (1992) and assume that agents ignore those actions which are not informative.

Definition 4 ((Un)Informative Action) *The action of an agent i, x_i, is (un)informative if that agent's signal, s_i, can (not) be inferred by observing x_i*

We can now describe the decision rule for each individual by employing an approach introduced by Acemoglu et al. (2011) which established that posterior beliefs can be expressed as an additive decomposition of *private beliefs* and *social beliefs*. An individual's private belief is defined as the probability $P(\omega = A \mid s_i)$ and their social belief as $P(\omega = A \mid O_i)$. The following result is a modification of Proposition 2 in Acemoglu et al. (2011) which accommodates our tie-breaking assumption and yields a type-dependent decision rule, where an individual's type is their private signal.

[8] This assumption is also made in Anderson and Holt (1997), and, while not entirely innocuous, it appears to have been supported by experimental findings that individuals have a tendency to overweight their private signal relative to information obtained through observation (Anderson and Holt, 1997; Nöth and Weber, 2003; Goeree et al., 2007; Weizsäcker, 2010). A common alternative assumption is to suppose that individuals randomize over their alternatives when indifferent (Bikhchandani et al. (1992); Acemoglu et al. (2011)).

Proposition 1 *The action of agent i, x_i, satisfies the following decision rule*

- *If $s_i = a$,*

$$x_i = \begin{cases} A, & \text{if } P(A \mid O_i) \geq 1 - q \\ B, & \text{if } P(A \mid O_i) < 1 - q \end{cases}$$

- *If $s_i = b$,*

$$x_i = \begin{cases} A, & \text{if } P(A \mid O_i) > q \\ B, & \text{if } P(A \mid O_i) \leq q \end{cases}$$

PROOF: All proofs are provided in online appendix B.

The form of the optimal decision rule highlights the important role that socially conveyed information plays in influencing the decisions of individuals. Additionally, we have discussed the possibility of information cascades in which social beliefs outweigh private beliefs to the extent that individuals simply disregard their private information. This characterization of the individual decision rule allows us to establish precise conditions under which an individual's action will be informative or uninformative. Specifically, we describe the following property which relates the informativeness of an individuals' action to their social belief.

Property 1 *The action of an agent i, x_i, is uninformative if $P(A \mid O_i) \in (0, 1-q) \cup (q, 1)$; otherwise, it is informative*

This property follows directly from the decision rule in Proposition 1 since, for extreme levels of social belief, an agent makes the same choice regardless of their private signal. To better understand how cascades and uninformative actions affect social learning in various network structures, we first consider what criteria must be met in order for a cascade to form. Specifically, we are interested in characterizing conditions under which an individual is susceptible to falling into a cascade. The following result demonstrates that an individual must have observed a minimum number of prior actions in order to possibly be in a cascade.

Lemma 1 *For any network G, in the binary states case, an individual i can be in a cascade only if $|O_i| \geq 2$*

Intuitively, an individual must have inferred enough socially conveyed information in order for it to be rational for them to disregard own private information. While having observed at least two prior actions is necessary for an individual to be in a cascade, it is not sufficient. It is possible for private signals to be distributed in such a way that an information cascade never occurs in a finite network. As an example, consider the complete history network, G^C and a sequence of private signals that continuously alternates between a

and b. In this scenario, every agent chooses the action corresponding to their private signal, a cascade never forms, and all actions are informative. Furthermore, it is possible, albeit less trivial, to construct a signal sequence in such a way as to prevent the formation of a cascade in more general sequential network structures.

It is useful at this point to define a property of network structures that will prove important going forward in our analysis. It reflects, from the perspective of the final decision-maker, the number of nodes within the network where the action is guaranteed to be informative. We say that a network is *k-informative* if $k < n$ actions coincide with private signals regardless of the distribution of signals.

Definition 5 (k-Informative Network) *A sequential network G is k-informative if the actions of k non-final agents are guaranteed to be informative for any sequence of private signals, $0<k<n$*

It is worth noting that, while this definition could be extended to allow for 0-informative networks, these never arise in the binary states case with equal priors that we consider here; however, in general, it is possible for a network to be 0-informative.[9]

In some networks, it is impossible for any non-final agent to fall into a cascade for any signal draws. This in turn guarantees that the action of every non-final agent is informative about their private signal and the network is $(n-1)$-$informative$. In these networks, all privately held information is aggregated, and the final decision maker will effectively have n private signals at their disposal when making a decision. We say that a network with this property is *fully informative.*

Definition 6 (Fully Informative Network) *A sequential network G is fully informative if the actions of all non-final agents are guaranteed to be informative for any sequence of private signals*

While the concept of k-informative and fully informative networks apply generally to observational learning environments, it will be useful for the present analysis to describe properties of these networks which are specific to the environment considered here. For the binary states case with equal priors, we describe the relationship between a k-informative network and the information structure it generates. Furthermore, since fully informative networks are simply $(n-1)$-informative, this relationship extends to these network structures as well. These properties are described formally below.

[9]Consider a binary states environment with asymmetric priors such that $P(A) = p_0 > 1/2$ and signal precision $P(a \mid A) = P(b \mid B) = q < p_0$. In this setting, each agent i will choose action $x_i = A$ regardless of their private signal, and therefore, all actions are uninformative. This case is outside the scope of the present study since in sequential networks that are 0-informative, no private information is aggregated, and no observational learning occurs.

Property 2 *A k-informative sequential network, G, generates an information structure* $I(G)$ *such that* $|O_i| \leq 1$ *for all* $i \in N_1$*,* $|O_i| \geq 2$ *for all* $i \in N_2$*, and* $O_n = \{x_1, x_2, \dots, x_{n-1}\}$*; where* N_1 *and* N_2 *are disjoint subsets of* N *such that* $N_1 \cup N_2 = \{1, 2, \dots, n-1\}$ *and* $|N_1| = k$

Property 3 *A fully informative sequential network, G, generates an information structure* $I(G)$ *such that* $|O_i| \leq 1$ *for all* $i \in \{1, 2, \dots, n-1\}$ *and* $O_n = \{x_1, x_2, \dots, x_{n-1}\}$

As previously discussed, the formation of a cascade limits the aggregation of private information because individual actions do not inform about private signals. Then, an important consideration for our comparison of network structures is the likelihood of cascade formation in various structures. We begin by considering the benchmark case of the complete history network, G^C, and demonstrate that cascades can only begin at specific points in the sequence of decision makers.

Proposition 2 *In the complete history network,* G^C*, an information cascade can start with an agent* $i \geq 3$ *and only if* i *is odd.*

From Lemma 1 and Proposition 2 we can infer that the number of agents with which a cascade can begin depends on the structure of the network. This raises a question about how more complex changes in the network, and, more specifically, changes in the number and location of these points at which cascades can begin may affect a network's susceptibility to cascade formation. This is particularly important in the present environment given that information cascades are not necessarily a permanent state. In the complete history network, Bikhchandani et al. (1992) note that a cascade, once formed, is never reversed. While there are many proofs of this result in the literature, this statement does not hold for more general network structures.

Consider, as a counter example, a network comprised of two separate chains of an equal number of individuals which eventually converge into a single chain. If every individual observes all prior actions in their own chain but none from the other chain, these sequences act as distinct complete history networks until they converge. It is possible that the chains will herd on different actions. Then, if we consider the choice of the first individual after which these isolated chains converge, this agent, observing the same number of informative actions on respective chains, will choose the action corresponding to own private signal, thereby ending both cascades.

3.3 Individual and Social Criteria

In this section we address the key question of this paper regarding which network structures will yield better outcomes or greater accuracy. To address this question, we propose two criteria to compare outcomes across networks, which are motivated by observations that while, for example, in some committees decisions are taken by voting and the overall accuracy of the group matters most, in other types of organizations, it is the final decision maker whose choice is most important.

Given the decision rule described in Proposition 1, for a fixed network structure and signal precision, the actions of all agents are determined solely by the sequence of private signals. Therefore, in order to compare networks based on outcomes, we consider expected payoffs over the set of possible signal sequences which may arise. Let $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ denote a vector of individual signals, $S(n)$ be the set of all possible signal sequences of length n, and $x_i(\mathbf{s})$ the choice of any agent i for a given signal sequence $\mathbf{s}$. Then, we define the expected payoff of each agent i as follows.

$$E_{S(n)}[u_i(x_i, \omega; G)] = \sum_{\mathbf{s} \in S(n)} P(\mathbf{s}) u_i(x_i(\mathbf{s}), \omega; G) \tag{1}$$

3.3.1 Individual Criterion

In some settings it may be most important that a particular, key, individual ultimately makes the correct decision. Consider, as an example, a business setting in which employees sequentially make recommendations regarding an investment opportunity to their immediate superiors. This process continues up the organizational chain until recommendations are made to the CEO who then makes the final decision of whether or not to invest. In this scenario, it is much more important that the final agent (the CEO) makes the correct decision than it is that any other individual gives a correct recommendation. Therefore, an analysis of the merits of various organizational structures should focus on the expected payoff of the final agent in the sequence.[10] Our *individual criterion* takes this approach by comparing networks (of equal size) based on the expected payoff of the final agent in the sequence.[11]

Following equation (1), we can express the expected payoff of the final agent.

$$E_{S(n)}[u_n(x_n, \omega; G)] = \sum_{\mathbf{s} \in S(n)} P(\mathbf{s}) u_n(x_n(\mathbf{s}), \omega; G) \tag{2}$$

[10] While they do not model the choice of the final decision maker explicitly, Jackson et al. (2019) takes a similar approach in examining whether it is possible for the final agent to learn the true state when messages, which may be randomly mutated or intentionally biased, are passed along chains of individuals which ultimately converge.

[11] While the individual criterion could be discussed in regards to any arbitrary agent in the sequence, analyzing the final agent is most meaningful in economic and social settings since many organizations have a chairperson or executive officer which makes the ultimate decision. The theory to analyze non-final agents is a straightforward extension of the analysis presented in this paper. Individuals observing no prior actions are covered by standard choice under uncertainty models. Other non-final agents are described by individual criterion considering a sub-network in which they are the final agent.

In what follows, we examine how the final agent's expected payoff differs under various network structures. We begin by noting that for very small groups all networks yield the same expected payoff for the final agent.

Proposition 3 *For fixed n and q, if $n \leq 3$, then the expected payoff of the final agent is the same under any possible network structure*

In such small network structures, an information cascade cannot start before the final agent, and all prior actions are guaranteed to be informative for every sequence of private signals. Thus, the final agent always has access to n private signals when making a decision and, regardless of the structure of the network, makes the same choice. More generally, this reasoning can be extended to show that, for any fixed number of agents, the expected payoff of the final agent is equal under all network structures that are fully informative. Furthermore, as shown in the next result, for a fixed number of agents, the expected payoff of the final agent is maximized by any fully informative network structure.

Proposition 4 *For fixed n and q, the expected payoff of the final agent is maximized by any network that is fully informative*

This demonstrates that the best outcome for the final agent could be achieved by every other member having a direct link to him. These are not, however, the kinds of network structures we might observe in real world organizations. In environments where the final agent has the ability to directly manipulate the structure of the network, why would she choose to implement an organizational structure that is not fully informative?

One possible explanation is that it may be exceedingly costly, in time and effort, for this individual to directly interact with every other person in the organization. We can explicitly examine this scenario by modifying our individual criterion to account for the costs of direct links to the final agent. Specifically, assume that the final agent must pay a fixed cost, $c > 0$, for each incoming link. A standard measure of individual connectivity is the *in-degree* of an agent which describes the number of other individuals that link directly that agent. In our framework, the in-degree of the final agent in a network G is defined as $\eta_n^{IN}(G) \equiv |\{j \in N : jn \in G\}|$. By Definition 1, the final agent must have $\eta_n^{IN}(G) \geq 1$ in all sequential networks. If direct links are costly, the final agent must then consider the cost of maintaining additional links relative to their expected benefits. This tradeoff is captured by the following expression of the final agent's utility, $\tilde{u}_n$, which is a modification of our earlier expression for the expected payoff of the final agent.

$$E_{S(n)}[\tilde{u}_n(x_n, \omega; G)] = \sum_{\mathbf{s} \in S(n)} P(\mathbf{s}) u_n(x_n(\mathbf{s}), \omega; G) - c\eta_n^{IN}(G) \tag{3}$$

When links are costless, Proposition 4 demonstrates that the expected payoff of the final agent is maximized by any network structure that is fully informative. In the presence of sufficiently costly direct links, there is a unique network structure that maximizes the expected payoff of the final agent. This network generates an information structure with $O_i = \varnothing$ for $i \in \{1, 2, \ldots, n-2\}$, $O_{n-1} = \{x_1, x_2, \ldots, x_{n-2}\}$, and $O_n = \{x_1, x_2, \ldots, x_{n-1}\}$, and therefore, is $(n-2)$-informative. This structure minimizes the expected loss over all networks which are not fully informative since at most one private signal cannot be inferred. Additionally, it minimizes the link costs to the final agent by having her interact directly with only one prior agent. Returning to our CEO story, we might think of this network as one where the CEO hires an assistant to speak with each employee directly and then advise the CEO on the best course of action.

3.3.2 Social Criterion

On the other hand we are interested in taking a standard welfare analysis approach to determine the network structure which maximizes aggregate expected payoffs. Indeed, a social planner would be interested in structuring a network that aggregates private information in such a way as to maximize the aggregate expected payoffs of all individuals. The overall accuracy of all members of the organization or committee is essential when decisions are taken by voting. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ denote the vector of all individual choices. Then, the *social criterion* compares network structures based on aggregate expected payoffs which can be expressed as follows:

$$E_{S(n)}[U(\mathbf{x}, \omega; G)] = \sum_{\mathbf{s} \in S(n)} P(\mathbf{s}) \sum_{i=1}^{n} u_i(x_i(\mathbf{s}), \omega; G) \tag{4}$$

Changes in the network structure, G, affect observational learning and individual choices through changes in the information structure, $I(G)$, since individuals observing different sets of prior actions face inherently different choice problems. Furthermore, the formation of a cascade, the likelihood of which differs under various structures, limits the amount of information some individuals are able to infer from observing past actions. To understand the welfare impacts of these various network effects, we begin by comparing the expected payoffs of individuals in different positions within a fixed network.

As previously indicated, individuals with $|O_i| \leq 1$ always choose the action that corresponds to their private signal and have expected payoff:

$$E_{S(n)}[u_i(x_i, \omega; G)] = q \left[\sum_{r=0}^{n-1} \binom{n-1}{r} q^r (1-q)^{n-1-r} \right] V, \quad \text{if } |O_i| \leq 1 \tag{5}$$

For individuals with $|O_i| = 2$, both actions observed are guaranteed to be informative for any sequence of private signals. In any event where these actions differ, the agent's social belief is $P(A \mid O_i) = 1/2$ and they choose the action corresponding to their private signal. However, if the two observed actions coincide, agent i falls into a cascade and imitates the action taken by the two observed agents. Therefore, we can express the expected payoff of any individual with $|O_i| = 2$ as follows:[12]

$$E_{S(n)}[u_i(x_i, \omega; G)] = \left[\left(q^2 + 2q^2(1-q)\right) \sum_{r=0}^{n-3} \binom{n-3}{r} q^r (1-q)^{n-3-r} \right] V, \quad \text{if } |O_i| = 2 \tag{6}$$

Direct comparison of Equations (5) and (6) reveals that, regardless of the structure of the network, if there is an individual who observes exactly two actions, he is strictly better off, in expectation, than individuals observing zero or one prior action.

Proposition 5 *For any network, G, and any agents i and j such that $|O_j| \leq 1$ and $|O_i| = 2$,*

$$E_{S(n)}[u_i(x_i, \omega; G)] > E_{S(n)}[u_j(x_j, \omega; G)]$$

Comparisons between two agents i and j such that $|O_i| > |O_j| = m$ for $m \geq 2$ are less straight forward. While it is true that individuals are better off with more information, observing additional actions does not always imply that more information is inferred since some of these actions may be uninformative. However, what we can say with certainty is that individuals are weakly better off than those whose actions they observe.[13] The underlying intuition of this result is that they are better off ex ante because, for every sequence of private signals, they must have at least as much information when making their decision.

Proposition 6 *For any network, G, and any two agents i and j such that $x_j \in O_i$,*

$$E_{S(n)}[u_i(x_i, \omega; G)] \geq E_{S(n)}[u_j(x_j, \omega; G)]$$

The intuition of this result implies that in the complete history network, beginning with the third agent, non-final agents are better off in expectation than non-final agents in a fully informative network. We can show that, for $n \geq 4$, these improvements outweigh the

[12] As shown in the proof of Proposition 5, Equation (5) can be reduced to qV and Equation (6) can be reduced to $[q^2 + 2q^2(1-q)]V$.

[13] Prior literature has described a similar *improvement principle* which details the improvement in an individual's ex ante expected payoff relative to one of their neighbors whose action they observe. For a comprehensive discussion of the improvement principle and its importance in the study of social learning in networks, see, for example, Acemoglu et al. (2011), Lobel and Sadler (2015), and Golub and Sadler (2016).

increase in expected payoff to the final agent, and therefore, aggregate expected payoffs are higher in the complete history network than in fully informative networks.

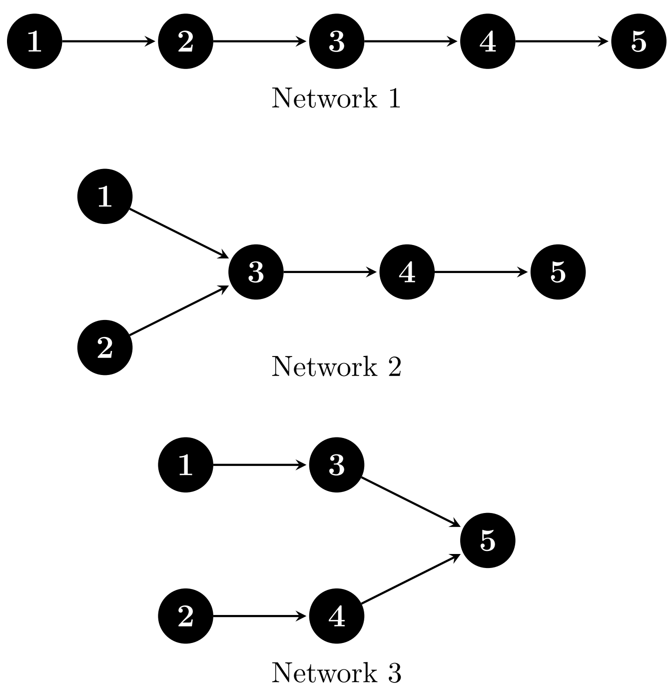
Proposition 7 *Aggregate expected payoffs are at least as large in the complete history network as in any fully informative network and strictly greater if* $n \geq 4$.

We therefore note the tradeoff between the expected payoff of the final agent and that of the group overall. A fully informative network provides the most independent private signals to the final agent; however this comes at the cost of this information being unavailable to any intermediate agents, thereby reducing their expected earnings.[14].

4 Experimental Design

In this section, we describe the design and procedures of our experiment aimed to test theoretical predictions. We employ two treatment variables. First, we vary the network structure, examining the three structures shown in Figure 1. Network 1 is the complete

Figure 1: Experimental Networks



history network and Network 3 is fully informative. We can therefore directly examine our predictions about the tradeoff between the payoff of the final agent and the aggregate group

[14]An illustration of this tradeoff is presented in Appendix A Figure 5

payoff. Network 2 is not the complete history network, but it is outcome equivalent to Network 1 under the tie-breaking assumption discussed earlier. The selected networks allow us to examine the effects of network structure and compare observed behavior with our theoretical results discussed in Section 3.

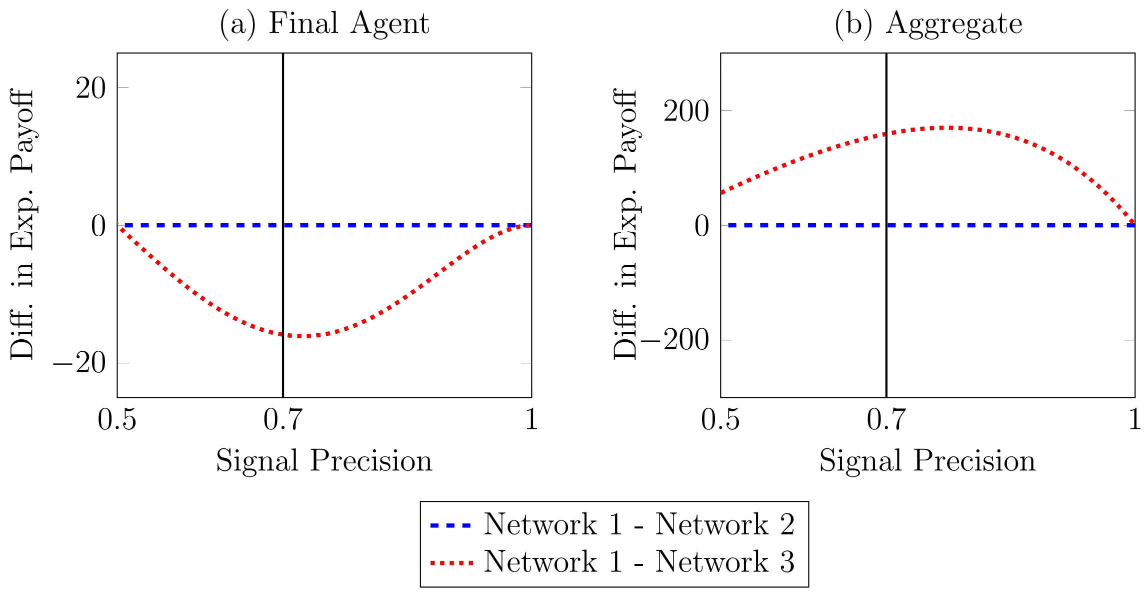
The second experimental treatment variable is the assignment method to positions within the network. As a control, subjects are assigned to positions randomly. A treatment variation here is the endogenous assignment which is implemented via an auction-type allocation mechanism developed for this setting. Although, unlike typical auctions, the number of bidders is equal to the number of positions, so everyone is guaranteed to be assigned, our mechanism allows us to elicit subjects' evaluations of the expected payoff premiums of being in later positions. Specifically, in a fully informative network, the final agent's expected payoff dominates any earlier agent's, and in the complete history network, expected payoff is weakly increasing in the number of observed actions. Endogenous assignment treatments, employing the allocation mechanism described in subsection 4.2, allow us to investigate subjects' beliefs manifested in their bidding behavior. The endogenous allocation stage treatments are added for Networks 1 and 3, our main networks of interest. In addition to investigating subjects' demand for information and anticipated benefits of occupying specific positions we are interested in the potential impact of heterogeneity.

4.1 Network Structures and Experimental Parameters

We selected networks comprised of five players because they are large enough for social learning to affect individual decisions (thus allowing us to test predictions) but small enough to avoid additional complexity. The specific networks chosen represent important structures discussed in theoretical analysis in Section 3. Network 1 is the baseline, complete history network, in which each individual observes the actions taken by everyone before them. Network 3 is a fully informative network structure where the final agent learns all five independent signals. Network 2, while distinct from Network 1, produces theoretically identical choices for each position within the network for every possible sequence of private signals. Thus, any differences in outcomes between Network 1 and Network 2 may indicate a behavioral effect arising from a change in the network structure, including non-Bayesian beliefs.

As described in Section 3, two possible events, A and B, are equally likely to occur. The payoff for making the correct choice is \$9.00. To determine the parameter value for signal precision q, we obtained theoretical expected payoffs across Networks 1-3. Since the signal is informative, the possible range of values of q is $(\frac{1}{2}, 1)$. Figure 2 shows the difference in the ex ante expected payoffs for the final agent (individual criterion) and the group (social criterion) as a function of the signal precision.[15]

Figure 2: Difference in Expected Payoffs by Signal Precision



Since Network 1 and Network 2 produce theoretically identical choice sequences, the difference in expected payoffs is zero in both cases. Comparing Network 1 and Network 3, we see that, as expected, Network 3 yields a higher payoff to the final agent, while Network 1 generates higher group payoff. For both criteria, the difference is maximized in the same neighborhood of q, and based on the graphs we selected the value $q = 0.7$.[16] While the actual value of difference maximizing signal precision is slightly higher in both cases, we round down for the following reasons. First, for subject understanding, it is simpler than presenting a more precise but complex number. Second, even greater than 0.7 signal precision will increase the likelihood of uninteresting cases – signal draws where actions become uninformative very early. Lastly, if signals are overly informative, subjects might be uninterested in social learning and ignore prior actions, which would make it more difficult for us to test our theory, including the consistency with the tie breaking assumption. Since our primary goal is to examine how changes in network structure affect social learning, 0.7 should provide a good environment and the right balance between above considerations. Overall, this choice of signal precision should give our experiment the best opportunity to

[15] For all figures, amounts are in cents and correspond to our experimental parameter of \$9 or 900 cents payoff for the correct choice.

[16] We assigned more weight to the optimal q for the difference in the payoffs of the final agent since testing that difference is a greater challenge. The magnitude of the difference is less than 20 cents which is unlikely to return statistical significance given random signal draws and possible deviations and heterogeneity of behavior. We could not increase the overall payoff given standard earning rates for the experiments at XS/FS experimental lab as well as budgetary considerations.

investigate treatment effects.

4.2 Endogenous Position Assignment

In our main treatments, positions within the networks are assigned randomly once subjects have been matched into groups of five. This corresponds to the standard convention in prior information cascade and herding literature which assumes that the sequence in which individuals act is exogenously determined.[17] For our endogenous assignment treatments, network positions are allocated via an auction allocation mechanism. At the beginning of each round, after new groups of five have been formed, subjects participate in an auction to determine their positions within the network. The auction mechanism was designed for our setting with the main objective that it reveals each subject's valuations of payoff premiums of all positions.

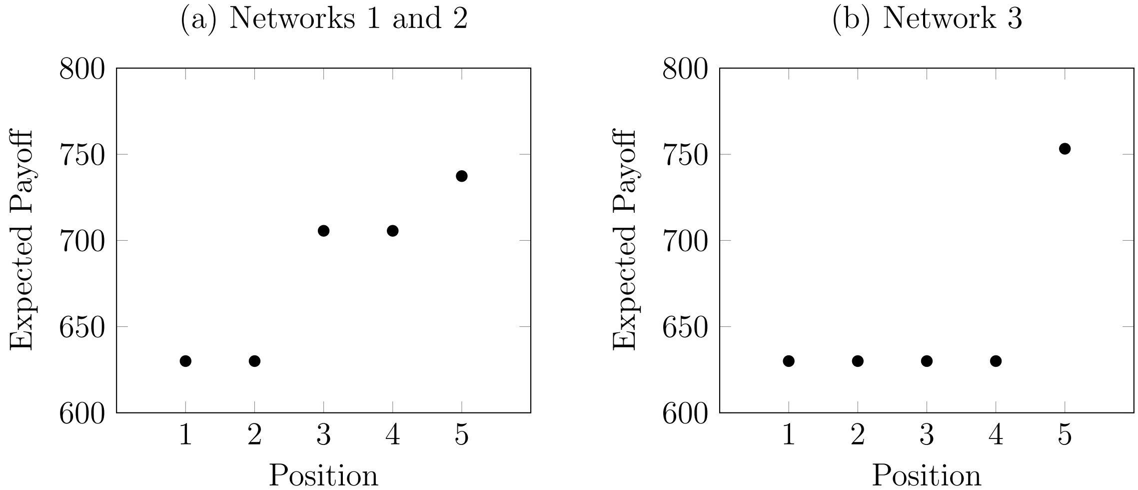
Subjects are given an endowment of 100 cents ($1) and asked to submit a bid between 0 and 100 cents for each of the five positions in the network structure. This results in a total of 25 bids submitted within each group, each corresponding to a unique subject-position pair. The highest among these is identified and that subject is assigned to that position. All bids for that position and also by that subject are then removed from further consideration. There are then 16 subject-position bids remaining. The highest among these is identified and that subject is assigned to that position, leaving 9 remaining bids. This process is repeated until all subjects have been assigned a position in the network. Subjects pay a price equal to the next highest bid, among all 25 of the originally submitted bids, for the position they are assigned. In the event that a subject submits the lowest bid for their assigned position, they pay a price of zero. Any remaining balance from their bidding endowment is added to their earnings for the round.

Examining subjects' bidding provides insights into the degree to which they understand the value of information and the benefit afforded to later decision makers. The ex ante expected payoffs for every network-position are shown in Figure 3. We see that there is a premium to acting later in the sequence since these individuals are able to infer additional information by observing earlier choices. We expect that bids are non-decreasing in position and outline risk neutral bidding below.

Unlike most auctions, our position allocation mechanism has the same number of bidders and positions, so every subject is guaranteed to be assigned a position within the network. Therefore the valuation of each position by risk neutral subjects is the premium that this po-

[17]Exogenous order of decisions is explicit in the original formulation of social learning models (Bikhchandani et al., 1992; Banerjee, 1992), and this assumption was subsequently adopted by experimental studies (e.g., Anderson and Holt (1997); Nöth and Weber (2003); Çelen and Kariv (2004a, 2005)).

Figure 3: Individual Expected Payoffs



sition is expected to generate compared to position(s) with no opportunity for observational learning. For example, in Network 1, the expected payoff is 630 in positions 1 and 2 (0 and 1 observed actions respectively), therefore the expected premium is 76 for positions 3 and 4, and it is 107 for position 5. In equilibrium subjects bid own valuations for the position, so agents will bid 0 for positions 1 and 2 and respective premiums for the later positions. In the theoretical benchmark case of homogeneous risk neutral agents, bids for all positions will be tied (broken randomly), effectively resulting in random assignment while allowing us to elicit valuations, which is our main objective. The assignment process among heterogeneous subjects during the experiment will reflect their likely heterogeneous valuations and bidding which may be affected by their risk preferences, understanding of the environment and the advantage of later positions, and other behavioral phenomena.

4.3 Treatments

Based on network structures and random versus endogenous position assignment, we conducted five treatments. The different treatments, as well as the number of sessions and subjects for each, are shown in Table 1. For the remainder of the paper we refer to network and position assignment conditions using the treatment abbreviations shown in Table 1.

Table 1: Experimental Treatments

Treatments	Sessions	Subjects
Network 1: Exogenous (N1)	2	40
Network 2: Exogenous (N2)	2	35
Network 3: Exogenous (N3)	2	40
Network 1: Endogenous (N1A)	3	50
Network 3: Endogenous (N3A)	3	50

4.4 Procedures

All experimental sessions were conducted at the XS/FS laboratory at Florida State University using z-Tree (Fischbacher, 2007). A total of 215 subjects participated in 12 sessions with each subject participating in only one session. All participants were recruited using ORSEE (Greiner, 2015) from a pool of FSU students who had previously registered to participate in experiments. Upon arrival, subjects were seated at private computer terminals. At the beginning of each session, instructions were distributed and then read aloud by the experimenter.[18] Additionally, prior to the beginning of the experiment, subjects completed a practice stage to familiarize themselves with the program interface. In endogenous assignment treatments subjects also received instructions and illustration of the allocation mechanism.

At the beginning of each round, subjects were matched into groups of five and assigned to positions within a network. Prior to making a decision, each subject was given their private signal that round, A or B, and knew that it corresponds to the true state with probability 0.7. Events and private signals were randomly generated prior to the experiment; summary statistics for the signal sequences are provided in Appendix C. Additionally, subjects were shown the choices made by others in their group that acted before them (if any) as specified by the network structure. Subjects made decisions in sequence, according to their position in the network, with their sole task being to make a prediction about which of these events had occurred.

A round ended after every subject has made a decision. Subjects were then told which of the two events actually occurred and informed about their earnings. Subjects earned \$9 for the round if their choice corresponded to the event and nothing otherwise.[19] Each experimental session consisted of seven decision making rounds. In each session, subjects interacted within the same network structure in all rounds, although their position within the network might have changed across rounds. Groups were randomly rematched before the start of each round.

After all seven rounds of the main part of the experiment had been completed, subjects participated in the (over)confidence elicitation instrument. The instrument was designed for this study to investigate whether subjects' consistency with Bayesian updating and propensity for observational learning are affected by their general confidence in their knowledge.[20]

[18] Sample experimental instructions are available in online Appendix C.

[19] The experimental findings of Anderson (2001) suggest that this is more than sufficient to induce subjects to approach the task seriously.

[20] Our instrument is an improvement on select existing confidence elicitation tasks as it allows measuring subjects' assessment of the precision of their knowledge and provides a continuous measure of (over)confidence. We find evidence that confidence affects decisions and consider using this data and re-

At the end of the experiment, one round was randomly selected for payment. Each session lasted between 60 and 90 minutes. On average, subjects earned \$14.48 including a \$7 show up payment.

5 Results

5.1 Consistency with Bayesian Choices

We start with noting the consistency of choices with theory. Since individual and group payoffs depend on randomly drawn signal sequences, we compare observed data to simulations of Bayesian agents. Thus, our hypotheses are whether subjects' decisions are consistent with actions of simulated Bayesian agents conditional on experimental signal draws. Table 2 shows that, overall, about 15% of choices are not consistent with the benchmark of all Bayesian agents. These deviations lead to efficiency loss which persists even under a more favorable evaluation paradigm: where we evaluate consistency with Bayesian choice based on experimental history. The loss of efficiency caused by the accumulation of "incorrect" choices accounts for only 0-4%, and the efficiency loss remains at 12%.

Table 2: Experimental Choices Consistent with Bayesian Updating

	N1	N2	N3	N1A	N3A	Total
All	233	204	250	291	309	1287
Bayesian	(0.8321)	(0.8327)	(0.8929)	(0.8314)	(0.8829)	(0.8551)
Conditional	242	215	250	304	318	1329
on History	(0.8643)	(0.8776)	(0.8929)	(0.8686)	(0.9086)	(0.8831)

Result 1 *About 15% of choices are not consistent with Bayesian predictions*

To investigate the causes of efficiency loss we first explore the tie-breaking assumption which states that, in the case of a 0.5 posterior, an agent follows own signal. Table 3 reports the frequency of such behavior.

We compare observed frequencies with two alternative benchmark frequencies: 0.5, which is consistent with the agent flipping a coin when posterior belief is 0.5, and 1, which indicates consistency with theoretical assumption. For all treatments, the observed frequency is significantly different from 1 ($p < 0.05$). At the same time, observed frequency is also significantly different than 0.5 ($p < 0.05$) in all except two treatments (N1: $p = .152$; and N2: $p = .138$). There is indeed a tendency to bias the tie breaking choice towards own

sults in a follow up paper.

Table 3: Experimental Choices Consistent with Tie Breaking Assumption

	N1	N2	N3	N1A	N3A	Total
Posterior 0.5	39	16	52	50	62	219
Consistent	24	11	41	33	44	153
Frequency	(0.6154)	(0.6875)	(0.7885)	(0.6600)	(0.7097)	(0.6986)

signal compared to a random draw, however the assumption of going with own signal is not supported.

Result 2 *The tie-breaking assumption is not supported by data*

A posterior of 0.5 can only occur for an odd number of observed informative actions. So for treatments N3 and N3A all relevant observations are only when $|O_i| = 1$ (positions 3 and 4), and, as indicated above, it is significantly different from both frequencies, 0.5 and 1, of going with own signal. Table 4, below, reports the frequency of breaking the tie in favor of own signal for treatments N1, N2, and N1A. For Network 1 (N1 and N1A) the 0.5 posterior can occur for positions 2, 4, and 5. For Network 2 (N2) it can occur only for positions 4 and 5. In all three of these treatments, position 5 needs to be treated with caution since there are only 2-3 relevant observations for each network as reported in Table 4. The results indicate that the propensity to break the tie in favor of own signal is decreasing in the number of observed actions.

Table 4: Choices Consistent with Tie Breaking Assumption by Number of Observed Actions

Treatment		Number of Observed Actions		
		1	3	4
N1	Posterior 0.5	21	15	3
	Consistent	18	6	0
	Frequency	(0.8571)	(0.4000)	(0.0000)
N2	Posterior 0.5	-	14	2
	Consistent	-	10	1
	Frequency	-	(0.7143)	(0.5000)
N1A	Posterior 0.5	27	21	2
	Consistent	23	10	0
	Frequency	(0.8519)	(0.4762)	(0.0000)

Result 3 *Observed propensity to break a tie in favor of own signal decreases with the number of observed actions*

For treatments N1 and N1A, when only one action is observed, the frequency of going with own signal is about 0.85 which is significantly different from 0.5 in both treatments ($p < 0.01$). Furthermore, this frequency is not significantly different from 1 in the N1A treatment. However, when three actions are observed, the likelihood of going with own signal is not significantly different from 0.5, and is, in fact, lower than 0.5; so, the propensity switches towards giving more weight to social learning and going against own signal. We find additional support for Result 3 by testing pairwise comparisons of frequencies of following own signal in case of 0.5 posterior. That frequency is significantly greater in cases when one action is observed than in cases when three or four actions are observed ($p < .01$; using both proportions and t-test). In all but one cases of 0.5 posterior when 4 actions are observed subjects went against own private signal. Pooling together N1 and N1A data also shows a (marginally) significant difference between positions 3 and 4 ($p < .06$; using both proportions and t-test).

5.2 Accuracy of choice

Inconsistency with Bayesian decision making manifests as a loss in payoffs. Table 5 reports the difference between observed payoffs and Bayesian payoffs for each treatment and the number of observed actions. With only three (not significant) exceptions, all numbers in the table have a negative sign. Furthermore, all numbers where social learning matters (number of observed actions at least two) have a negative sign and some are statistically significant as indicated in the table. Payoff losses become significant for network positions with two observable actions, which is the first position where a subject can disregard own private signal based on observing the actions of others. We therefore conclude that the complexity of social information aggregation leads to payoff loss.

Result 4 *There is, on average, payoff loss compared to Bayesian benchmark. This loss is more pronounced when social beliefs can dominate private beliefs.*

Table 10, in Appendix A, confirms that the payoff loss is consistent with deviations from Bayesian updating. Furthermore, Table 5 indicates that payoff loss is pronounced for the final agent. The largest and most significant deviation is observed for the final agent in N3A. As explained in more detail in the next section, with endogenous assignment, subjects who are better at understanding the advantage of the final position (recall that N3 is the most favorable to the final agent) self select to be in the final position. Earlier positions end up being filled by subjects with less understanding and greater noise in decisions. We do not find significant differences in the payoff of the final agent between treatments. The closest

Table 5: Difference in Earnings (Experimental minus All Bayesian)

Treatment	Number of Observed Actions				
	0	1	2	3	4
N1	16.07	16.07	-80.36†	-48.21	-80.36†
	(27.44)	(36.65)	(51.59)	(58.14)	(47.59)
N2	-9.18	-	-91.84	-73.47	-128.57**
	(30.90)	-	(63.24)	(51.54)	(55.19)
N3	-40.18†	8.04	-	-	-64.29
	(25.80)	(31.42)	-	-	(44.49)
N1A	-12.86	-38.57	-102.86*	-12.86	-51.43
	(34.07)	(39.54)	(52.91)	(50.29)	(48.84)
N3A	-12.86	-6.43	-	-	-154.29***
	(18.27)	(26.83)	-	-	(40.63)

Average, robust standard errors (clustered at subject level) reported in parentheses

† $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

to the standard significance level difference ($p < 0.1$) is between treatments N3 and N3A indicting the endogenous selection effect.

Result 5 *Observed accuracy of choices of the final agent in networks N3 and N3A is not greater than in other networks for both Bayesian and experimental data.*

The hypothesis of higher payoff to the final agent in N3 and N3A compared to other networks is among most challenging to evaluate experimentally since the theoretical expected difference in payoffs is only 15.88 cents. Given random signal draws, even decisions of Bayesian agents do not result in manifestation of ex ante expected differences in payoffs as can be seen from Table 12 in Appendix A. Therefore, it is not that this result is not supported by data, but our data, given random signal draws realizations, does not allow us to statistically test it. An experimental design aimed specifically at studying the payoffs of the final agents may employ larger networks and the parameters resulting in greater payoff differences to give theory a statistical chance to be tested.

We can, however, use our setting to test the underlying property determining the performance of the final agent. Table 6 presents two panels that report the effect of the number of observed informative actions on individual earnings. The upper panel is based on simulated data for Bayesian agents and serves as hypotheses for the lower panel that reports experimental data. Table 6 indicates that, except for treatments with endogenous assignments, the payoff is increasing in the number of observed informative actions as predicted by theory. In the endogenous assignment treatments early choice inconsistencies erode the payoff to the final agent.

Table 6: Effect of Observed Informative Actions on Earnings

(a) All Bayesian

	N1	N2	N3	N1A	N3A
Number of Informative Actions Observed	19.87	19.30	36.01**	26.55†	38.57***
	(18.47)	(18.40)	(16.19)	(16.68)	(14.55)
Constant	673.10***	654.70***	612.50***	645.70***	604.30***
	(38.61)	(36.74)	(30.72)	(35.29)	(27.60)

(b) Experimental

	N1	N2	N3	N1A	N3A
Number of Informative Actions Observed	-23.08	-17.21	27.98**	8.15	1.67
	(21.00)	(23.07)	(13.11)	(18.00)	(15.19)
Constant	711.30***	646.30***	596.40***	633.90***	610.00***
	(42.71)	(41.92)	(30.15)	(38.28)	(24.96)

Robust standard errors (clustered at subject level) reported in parentheses

† $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We next investigate the accuracy of decisions of the group. In this regard, Networks 1 and 2 (treatments N1(A) and N2) are predicted to perform better. Figure 4 shows the distributions of the number of correct choices in a group across treatments. Despite random signal draws, distributions for All Bayesian agents reflect theory: distributions for N1, N2 and N1A are skewed to the right compared to N3 and N3A. Using correct unanimity as the measure of group success, for Bayesian agents, the frequency is expected to be about 0.5 for N1(A) and N2 and only about 0.23 for N3(A). Although experimental data indicates lower accuracy overall, the frequency of correct unanimity is greater in treatments N1(A) and N2 compared to N3(A).

Formal tests of network comparison are presented in Table 7. The upper panel of the table indicates the differences between treatments in proportions of groups that were unanimously correct using decisions of All Bayesian agents. As predicted by theory N1 and N2 perform better than N3 (including with endogenous entry). The lower panel of the table is based on experimental data and shows that, with the exception of comparisons for N3A, theoretical predictions are supported. These results confirm that network properties outlined in Section 3 lead to better group performance. We next turn to a more detailed analysis of networks with endogenous position assignment.

Result 6 *Theoretical predictions for group performance are supported by data, except for treatment N3A. Specifically, groups in treatments N1(A) and N2 are more accurate than those in N3.*

Table 7: Difference in Proportion of Groups with Five Correct Choices

(a) All Bayesian

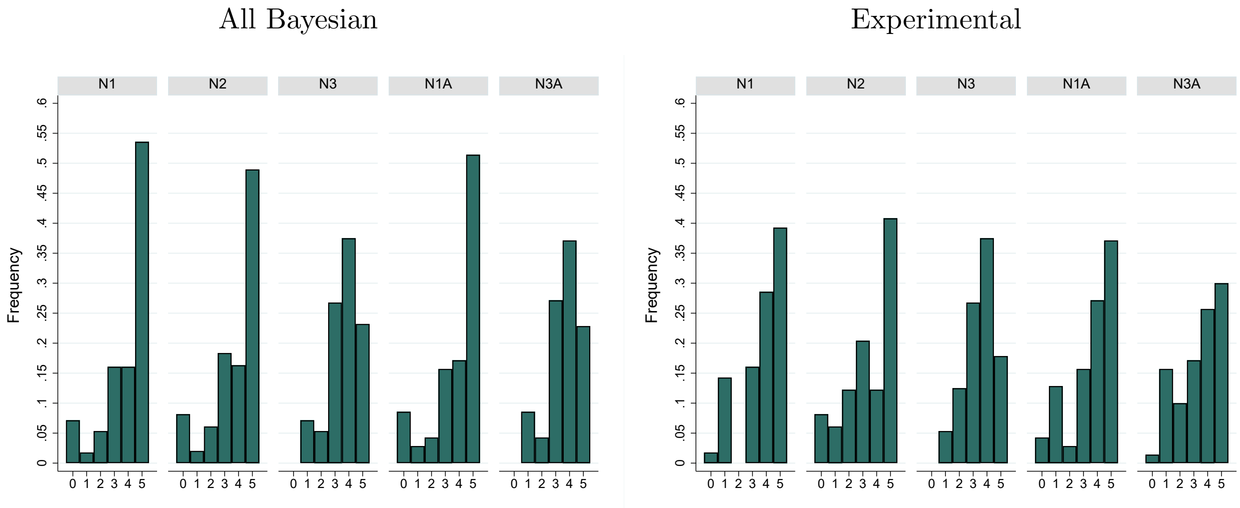
	N2	N3	N1A	N3A
N1	0.0459	0.3036***	0.0214	0.3071***
	(0.0977)	(0.0873)	(0.0895)	(0.0834)
N2	-	0.2577***	-0.0245	0.2612***
	-	(0.0910)	(0.0931)	(0.0873)
N3	-	-	-0.2821***	0.0036
	-	-	(0.0822)	(0.0755)
N1A	-	-	-	0.2857***
	-	-	-	(0.0780)

(b) Experimental

	N2	N3	N1A	N3A
N1	-0.0153	0.2143**	0.0214	0.0929
	(0.0959)	(0.0829)	(0.0871)	(0.0852)
N2	-	0.2296***	0.0367	0.1082
	-	(0.0869)	(0.0909)	(0.0891)
N3	-	-	-0.1929**	-0.1214†
	-	-	(0.0772)	(0.0750)
N1A	-	-	-	0.0714
	-	-	-	(0.0796)

Cells denote difference in proportions of row minus column
Standard errors reported in parentheses
† $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Number of Correct Choices in a Group



5.3 Demand for Information (Positions within Network)

Providing subjects with the mechanism to select into specific positions resulted in changed assignments compared to exogenous random method. The average number of different positions held by a subject during a session under exogenous assignment method was higher than in endogenous treatments (3.90 and 3.03, respectively). Table 8 reports the distribution of the number of different positions held by subjects across all rounds. With endogenous assignment the number of different positions occupied by a subject within a session decreased ($p < 0.01$) indicating heterogeneity in bidding and selection.

Table 8: Number of Different Positions (by subject within session)

	1	2	3	4	5
Exogenous	0	3	28	61	23
	(0.0000)	(0.0261)	(0.2435)	(0.5304)	(0.2000)
Endogenous	10	17	39	28	6
	(0.1000)	(0.1700)	(0.3900)	(0.2800)	(0.0600)

While there are no subjects who experienced only one position in exogenous treatments, the number of individuals who ended up in the same position in all decision rounds is positive in endogenous assignment treatments. Some noise in allocations in endogenous assignment treatments is expected and caused by occasional ties among bidders for later positions, consistently with theory, as well as bidding noise. Generally, in the endogenous assignment treatments we see the frequencies skewed towards a less diverse position distribution. Overall, this indicates heterogeneity in competitiveness for positions and recognition of the benefits of social learning on behalf of the majority of subjects.

Result 7 *There is evidence of (heterogenous) demand for information in endogenous positions assignment.*

Table 9 reports regression analysis of subjects' bids submitted for the network positions. Consistent with theory, subjects show demand for observational learning. Bid amounts increase with the size of the observation set. The structure of the network affects the bids, which are slightly greater for N3A. The benefit to the final agent in N3A is recognized by subjects ($p < 0.1$).[21] There is a slight tendency by subjects to pay premium for observing only one action. Recall that in case of 0.5 posterior, subjects' decision propensity changes with the number of observed actions. Subjects were significantly more likely to follow their signal when only 1 action was observed, while they effectively flipped a coin (or even sided with social beliefs against own signal) when facing 0.5 posterior after observing three or four actions. These decisions show that there is demand for information even if in the end subjects do not take it fully into account.

Result 8 *Subjects increase their bids for positions with the number of observed actions greater than 1. Bid premiums are consistent with expected payoff gains.*

6 Conclusion

We present a theoretical framework of choice under uncertainty where, in addition to own private signal, individuals observe previous choices of others. Our setting applies to situations when the decision is made once, thus learning from repetition is not possible, and we use directed networks to specify the observation sets of choices made by others. Observational learning in this setting relies on determining whether an observed action is informative and reveals an independent signal about the true state. We evaluate network performance using two criteria, individual (final agent) and social (group), and show that there is a tradeoff: properties that enhance performance under one criterion reduce the performance of the other criterion. For example, the complete history network, where all agents make decisions in a sequence and observe all prior actions, leads to higher social payoff. Conversely, a fully

[21]There is a positive premium that subjects are willing to pay to be the final agent in N3A compared to N1A albeit under lower confidence level. The results of both regressions when errors are not clustered on the subject level bring the significance level of extra premium of bids in N3A for the final position to $p < 0.01$. There is significant heterogeneity in recognizing the benefit of being the final agent in N3A. A subset of subjects who recognize it bid consistent with theory, while there is considerable noise and lack of consistency for other subjects.

Table 9: Subject Bids by Number of Observed Actions

	OLS	Tobit
N3A	2.96	7.44
	(2.61)	(7.49)
1 Observed Action	$0.73^{\dagger}$	5.03^{*}
	(0.45)	(2.68)
2 Observed Actions	4.51^{***}	17.72^{***}
	(1.43)	(4.92)
3 Observed Actions	18.78^{***}	36.75^{***}
	(4.52)	(8.62)
4 Observed Actions	20.35^{***}	37.21^{***}
	(4.74)	(9.11)
N3A × 1 Observed Action	3.68^{**}	5.14
	(1.47)	(4.08)
N3A × 4 Observed Actions	13.28^{*}	15.31
	(7.10)	(12.12)
Constant	4.97^{***}	-19.20^{***}
	(1.08)	(6.34)

Robust standard errors (clustered at subject level) reported in parentheses

† $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

informative network, where all agents are directly connected to the final agent, maximizes the payoff to the final agent at the expense of group payoff.

We design an experiment to test theoretical predictions using three network structures. We compare observed behavior to the benchmark of the performance of Bayes-rational agents and also investigate differences between treatments. In all treatments there is efficiency loss compared to Bayesian agents. One of the causes for deviations from Bayesian choices is violation of a tie breaking assumption in the case of equal posterior beliefs. The data reveals that this common assumption is not supported. Further, we find that the propensity to make decisions under equal posteriors changes with the size of the observation set: as the number of observed actions increases agents are more likely to go against own private signal. This result has important implications for theoretical literature on individual decision making with observational learning. This phenomenon suggests that if a strong consensus has not been reached by a number of agents, subsequent decision makers may disregard their private information and effectively erase the impact of new independent signals for future population.

We find support for our theoretical predictions regarding the relationship between network structure and social criterion. Networks which perform better under the social criterion tend to yield more accurate outcomes for groups in the experiment. Due to small magnitudes of payoffs and randomness of signal draws, we do not observe significant differences in the performance of final agents between treatments for both, experimental and Bayesian-simulated data. Despite this, we do find that the underlying property of the final agent's performance is supported by the data: accuracy of choice increases with the number of informative actions observed.

In settings where participants can self select to positions within the network, more sophisticated agents, who understand that later positions provide greater opportunities for observational learning, are generally more successful in getting later positions. This pushes other agents, who are more prone to make mistakes, to earlier positions within the network. Greater noise in decisions of earlier agents in turn erodes the benefit of observational learning and reduces the accuracy of the final agent in networks with endogenous position assignment. Examples where this phenomenon should be considered include multi-stage selection processes, such as hiring decisions that have to be approved by several levels of personnel. If the most sophisticated decision maker reaches the position of the CEO of the company, a talented potential worker may not be recognized by decision makers of the early stages of the selection process and ranked lower. Considering the propensity to disregard own signal with the number of observed actions as described above, the CEO might just go with the rankings received, which further highlights the importance of this phenomenon in organization design for optimal information aggregation.

Our results have implications for a range of social and economic settings; in particular, committee design and decision rules. Depending on whether a choice is set by the group (the strongest criteria being unanimous voting where every member has veto power) or by a single decision maker, our study informs on theoretical properties as well as behavioral consistencies that can be used in the design. The same properties of networks are also useful to inform how one might collect or present information to the decision makers. In the surgical procedure selection example described earlier, a person should try to talk to independent decision makers who made their choices without sharing information with each other. Moreover, if a person is misinformed about the structure of the network, for example if previous choices of others were part of a complete history network or generally were not part of a fully informative network, but are presented as the latter to the final agent, that would reduce the accuracy of the final agent, and made him side with the choices of others more than it is optimal. Future studies may also address settings where the final agent is uncertain about the structure of the network.

References

Acemoglu, D., M. A. Dahleh, I. Lobel, and A. Ozdaglar (2011). Bayesian learning in social networks. *The Review of Economic Studies 78*(4), 1201–1236.

Anderson, L. R. (2001). Payoff effects in information cascade experiments. *Economic Inquiry 39*(4), 609–615.

Anderson, L. R. and C. A. Holt (1997). Information cascades in the laboratory. *The American Economic Review*, 847–862.

Arieli, I. and M. Mueller-Frank (2018). Multidimensional social learning. *The Review of Economic Studies 86*(3), 913–940.

Banerjee, A. and D. Fudenberg (2004). Word-of-mouth learning. *Games and Economic Behavior 46*(1), 1–22.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics 107*(3), 797–817.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy 100*(5), 992–1026.

Çelen, B. and S. Kariv (2004a). Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review 94*(3), 484–498.

Çelen, B. and S. Kariv (2004b). Observational learning under imperfect information. *Games and Economic Behavior 47*(1), 72–86.

Çelen, B. and S. Kariv (2005). An experimental test of observational learning under imperfect information. *Economic Theory 26*(3), 677–699.

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2019). Testing models of social learning on networks: Evidence from two experiments. *Econometrica*.

Choi, S., D. Gale, and S. Kariv (2005). Behavioral aspects of learning in social networks: an experimental study. In *Experimental and Behavorial Economics*, pp. 25–61. Emerald Group Publishing Limited.

Choi, S., D. Gale, and S. Kariv (2012). Social learning in networks: a quantal response equilibrium analysis of experimental data. *Review of Economic Design 16*(2-3), 135–157.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Gale, D. and S. Kariv (2003). Bayesian learning in social networks. *Games and Economic Behavior 45*(2), 329–346.

Goeree, J. K., T. R. Palfrey, B. W. Rogers, and R. D. McKelvey (2007). Self-correcting information cascades. *The Review of Economic Studies 74*(3), 733–762.

Golub, B. and M. O. Jackson (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics 2*(1), 112–49.

Golub, B. and M. O. Jackson (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics 127*(3), 1287–1338.

Golub, B. and E. Sadler (2016). Learning in social networks. In Y. Bramoullé and A. Galeotti (Eds.), *The Oxford Handbook of the Economics of Networks.*

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association 1*(1), 114–125.

Grimm, V. and F. Mengel (2020). Experiments on belief formation in networks. *Journal of the European Economic Association 18*(1), 49–82.

Jackson, M. O., S. Malladi, and D. McAdams (2019). Learning through the grapvevine: the impact of message mutation, transmission failure, and deliberate bias. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3269543`.

Katz, M. L. and C. Shapiro (1985). Network externalities, competition, and compatibility. *American Economic Review 75*(3), 424–440.

Kübler, D. and G. Weizsäcker (2004). Limited depth of reasoning and failure of cascade formation in the laboratory. *Review of Economic Studies 71*(2), 425–441.

Lobel, I. and E. Sadler (2015). Information diffusion in networks through social learning. *Theoretical Economics 10*(3), 807–851.

Mossel, E., A. Sly, and O. Tamuz (2015). Strategic learning and the topology of social networks. *Econometrica 83*(5), 1755–1794.

Nöth, M. and M. Weber (2003). Information aggregation with random ordering: cascades and overconfidence. *The Economic Journal 113*(484), 166–189.

Rosenberg, D. and N. Vieille (2019). On the efficiency of social learning. *Econometrica 87*(6), 2141–2168.

Smith, L. and P. Sørensen (2000). Pathological outcomes of observational learning. *Econometrica 68*(2), 371–398.

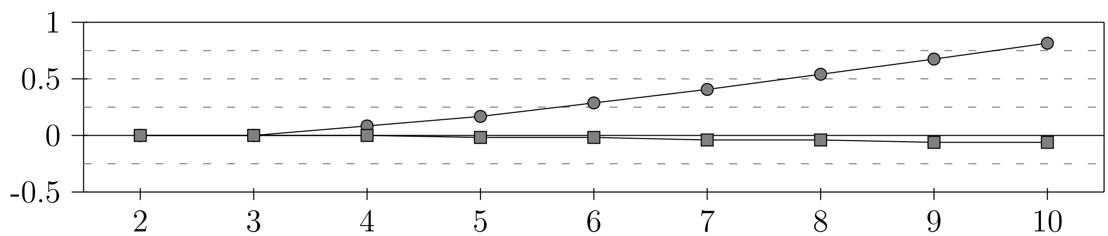
Thaler, R. H. (2016). Behavioral economics: past, present, and future. *American Economic Review 106*(7), 1577–1600.

Weizsäcker, G. (2010). Do we follow others when we should? A simple test of rational expectations. *American Economic Review 100*(5), 2340–60.

A Additional Tables and Figures

A.1 Social Criterion

Figure 5: Aggregate Payoff Difference



Horizontal axis indicates the number of agents (n). Circles indicate the aggregate expected payoffs of all non-final agents in the complete history network minus the aggregate expected payoffs of all non-final agents in a fully informative network. Squares indicate the difference in the expected payoff of the final agent in the complete history network minus the expected payoff of the final agent in the fully informative network. All calculations assume equally likely states, signal precision $q = 0.7$, and $V = 1$ for correct choice.

A.2 Supplementary Tables

Table 10: Frequency Against All Bayesian

Treatment	Number of Observed Actions				
	0	1	2	3	4
N1	3	5	11	17	11
	(0.0536)	(0.0893)	(0.1964)	(0.3036)	(0.1964)
N2	11	-	9	10	11
	(0.1122)	-	(0.1837)	(0.2041)	(0.2245)
N3	11	13	-	-	6
	(0.0982)	(0.1161)	-	-	(0.1071)
N1A	7	7	16	17	12
	(0.1000)	(0.1000)	(0.2286)	(0.2429)	(0.1714)
N3A	8	21	-	-	12
	(0.0571)	(0.1500)	-	-	(0.1714)

Table 11: Frequency Against Bayesian (Conditional on Experimental History)

Treatment	Number of Observed Actions				
	0	1	2	3	4
N1	3	5	10	13	7
	(0.0536)	(0.0893)	(0.1786)	(0.2321)	(0.1250)
N2	11	-	5	9	5
	(0.1122)	-	(0.1020)	(0.1837)	(0.1020)
N3	11	13	-	-	6
	(0.0982)	(0.1161)	-	-	(0.1071)
N1A	7	7	13	14	5
	(0.1000)	(0.1000)	(0.1857)	(0.2000)	(0.0714)
N3A	8	21	-	-	3
	(0.0571)	(0.1500)	-	-	(0.0429)

Table 12: Difference in Proportion of Final Agents making Correct Choice

(a) All Bayesian

	N2	N3	N1A	N3A
N1	0.0204	-0.0179	0.0143	-0.0143
	(0.0705)	(0.0643)	(0.0639)	(0.0615)
N2	-	-0.0383	-0.0061	-0.0347
	-	(0.0689)	(0.0684)	(0.0662)
N3	-	-	0.0321	0.0036
	-	-	(0.0620)	(0.0596)
N1A	-	-	-	-0.0286
	-	-	-	(0.0591)

(b) Experimental

	N2	N3	N1A	N3A
N1	0.0740	-0.0357	-0.0179	0.0679
	(0.0867)	(0.0775)	(0.0748)	(0.0786)
N2	-	-0.1097	-0.0918	-0.0061
	-	(0.0846)	(0.0821)	(0.0856)
N3	-	-	0.0179	0.1036
	-	-	(0.0723)	(0.0763)
N1A	-	-	-	0.0857
	-	-	-	(0.0735)

Cells denote difference in proportions of row minus column

Standard errors reported in parentheses

† $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Online Appendix: Proofs

Proof of Proposition 1

We begin by showing that the following condition holds

$$P(A \mid I_i) \geq \frac{1}{2} \iff P(A \mid s_i) + P(A \mid O_i) \geq 1$$

By Bayes' rule and our assumption that both states are equally likely, we have

$$P(A \mid I_i) = \frac{P(I_i \mid A)P(A)}{P(I_i \mid A)P(A) + P(I_i \mid B)P(B)} = \frac{P(I_i \mid A)}{P(I_i \mid A) + P(I_i \mid B)}$$

Thus, $P(A \mid I_i) \geq 1/2$ if and only if $P(I_i \mid A) \geq P(I_i \mid B)$. Conditional on the true state, ω, private signals and observed actions are independent. Therefore, for all $z \in \Omega$, we have

$$P(I_i \mid z) = P(s_i \mid z)P(O_i \mid z)$$

It then follows from the preceding relations that

$$\begin{aligned}
P(A \mid I_i) \geq \frac{1}{2} &\iff P(s_i \mid A)P(O_i \mid A) \geq P(s_i \mid B)P(O_i \mid B) \\
&\iff \frac{P(O_i \mid A)}{P(O_i \mid A) + P(O_i \mid B)} \geq \frac{P(s_i \mid B)}{P(s_i \mid A) + P(s_i \mid B)} \\
&\iff \frac{P(O_i \mid A)P(A)}{P(O_i \mid A)P(A) + P(O_i \mid B)P(B)} \geq \frac{P(s_i \mid B)P(B)}{P(s_i \mid A)P(A) + P(s_i \mid B)P(B)}
\end{aligned}$$

Applying Bayes' rule to each side of the previous inequality, we have

$$\begin{aligned}
P(A \mid I_i) \geq \frac{1}{2} &\iff P(A \mid O_i) \geq P(B \mid s_i) \\
&\iff P(A \mid s_i) + P(A \mid O_i) \geq 1
\end{aligned}$$

Suppose agent i has private signal $s_i = a$. Then, we have

$$\begin{aligned}
P(A \mid O_i) \geq 1 - q &\iff P(A \mid I_i) \geq \frac{1}{2} \implies x_i = A \\
P(A \mid O_i) < 1 - q &\iff P(A \mid I_i) < \frac{1}{2} \implies x_i = B
\end{aligned}$$

Suppose agent i has private signal $s_i = b$. Then, we have

$$P(A \mid O_i) \leq q \iff P(A \mid I_i) \leq \frac{1}{2} \implies x_i = B$$
$$P(A \mid O_i) > q \iff P(A \mid I_i) > \frac{1}{2} \implies x_i = A$$

The previous relations yield the stated decision rule. $\square$

Proof of Lemma 1

Here, we prove the contrapositive that an individual i with $|O_i| < 2$ cannot possible be in a cascade. Such an individual observes at most one prior action and, therefore, has $O_i \in \{\{A\}, \{B\}, \varnothing\}$. Their social belief, $P(A \mid O_i)$, is equal to q if $O_i = \{A\}$, $1-q$ if $O_i = \{B\}$, or $1/2$ if $O_i = \varnothing$. In any case, the decision rule derived in Proposition 1 dictates that they choose the action corresponding to their private signal. Therefore, there is no case in which an individual with $|O_i| < 2$ falls into a cascade. $\square$

Proof of Proposition 2

Let $G = G^C$ and suppose a cascade has not yet started (i.e., all prior actions are informative). By Lemma 1, an information cascade can start in this network structure no earlier than the third agent. Then, consider any agent $j \geq 4$ with j being even. It is easily verified that if agent $j-1$ (which is, by definition, odd) is not in a cascade their social belief must be $P(A \mid O_{j-1}) = 1/2$. It follows then that the social belief of agent j is such that $1 - q \leq P(A \mid O_j) \leq q$ and a cascade cannot possibly start with agent j. Therefore, in the complete history network, G^C, a cascade can start with an agent i only if i is odd. $\square$

Proof of Proposition 3

Fix q and note that $n \geq 2$ by construction. In the case of $n = 2$, there is a unique network structure that generates an information structure in which $O_1 = \varnothing$ and $O_2 = \{x_1\}$. For every signal sequence, the final agent chooses the action that corresponds to their private signal. In the case of $n = 3$, there is no longer a unique network structure. However, by Lemma 1, in any network with $n = 3$, an information cascade can start no earlier than with the final agent. Therefore, it is guaranteed that x_1 and x_2 are informative about their private signal and the final agent effectively has three independent signals when making their decision. Therefore, regardless of the structure of the network, the final agent makes the same choice for each realization of private signals. $\square$

Proof of Proposition 4

Fix n and q, set $V = 1$, let $\hat{G}$ denote a fully informative network structure, and let G denote an arbitrary network structure which is not fully informative. Let x_n and $\hat{x}_n$ denote the choice of the final agent in networks G and $\hat{G}$ respectively. By Definition 6 and Property 3, in G there must be at least one agent $j \in N \setminus \{n\}$ such that $|O_j| \geq 2$. Then, there must exist a signal sequence for which the action of agent j, x_j, is uninformative.

Consider a signal sequence for which the actions of $m < n-1$ non-final agents are informative in G (i.e., $n-1-m > 0$ actions are uninformative).[22] Let O_i and $\hat{O}_i$ denote the set of actions observed by agent i in G and $\hat{G}$ respectively. In $\hat{G}$ the final agent infers an additional $n-1-m$ signals by observing the actions of other agents. For any sequence of private signals, we define the following values, which describe the composition of information available to the final agent in each network.

$$\alpha_n = |\{x_\ell \in O_n : x_\ell = A \text{ and } P(A \mid O_\ell) \in [1-q, q]\}|$$
$$\hat{\alpha}_n = |\{x_\ell \in \hat{O}_n : x_\ell = A \text{ and } P(A \mid \hat{O}_\ell) \in [1-q, q]\}| - \alpha_n$$

The social belief of the final agent in network $\hat{G}$ can then be expressed as follows

$$P(A \mid \hat{O}_n) = \frac{q^{\alpha_n+\hat{\alpha}_n}(1-q)^{n-1-\alpha_n-\hat{\alpha}_n}}{q^{\alpha_n+\hat{\alpha}_n}(1-q)^{n-1-\alpha_n-\hat{\alpha}_n} + q^{n-1-\alpha_n-\hat{\alpha}_n}(1-q)^{\alpha_n+\hat{\alpha}_n}}$$

Applying the decision rule derived in Proposition 1, the choice of the final agent in $\hat{G}$ satisfies:

- If $s_n = a$,

$$\hat{x}_n = \begin{cases} A, & \text{if } \hat{\alpha}_n \geq \frac{n}{2} - 1 - \alpha_n \\ B, & \text{if } \hat{\alpha}_n < \frac{n}{2} - 1 - \alpha_n \end{cases}$$

- If $s_n = b$,

$$\hat{x}_n = \begin{cases} A, & \text{if } \hat{\alpha}_n > \frac{n}{2} - \alpha_n \\ B, & \text{if } \hat{\alpha}_n \leq \frac{n}{2} - \alpha_n \end{cases}$$

Define the following, where, by construction, $\hat{\alpha} \leq n-1-m$ must also hold.

$$Z_a^A = \left\{\hat{\alpha}_n : \hat{\alpha}_n \geq \frac{n}{2} - 1 - \alpha_n\right\} \qquad Z_b^A = \left\{\hat{\alpha}_n : \hat{\alpha}_n > \frac{n}{2} - \alpha_n\right\}$$
$$Z_a^B = \left\{\hat{\alpha}_n : \hat{\alpha}_n < \frac{n}{2} - 1 - \alpha_n\right\} \qquad Z_b^B = \left\{\hat{\alpha}_n : \hat{\alpha}_n \leq \frac{n}{2} - \alpha_n\right\}$$

There are now multiple cases to consider. In each we will show that the expected gains (EG) are at least as large as the expected losses (EL) for the final agent in $\hat{G}$ relative to G.

[22] Note that we need not consider any signal sequences for which the actions of all non-final agents are informative since it is necessarily the case that $x_n = \hat{x}_n$ for any such sequence.

Case 1 Suppose $P(A \mid O_n) \in (0, 1-q)$. For any $s_n \in \{a, b\}$, $x_n = B$ and the final agent in G makes the correct choice only if $\omega = B$. Therefore, the final agent experiences a gain in $\hat{G}$ if $\omega = A$ and $\hat{x}_n = A$. Conversely, they experience a loss if $\omega = B$ and $\hat{x}_n = A$. Thus,

$$EG = \frac{q^{\alpha_n}(1-q)^{m-\alpha_n}}{2}\left[q \sum_{\alpha \in Z_a^A} \binom{n-1-m}{\alpha} q^{\alpha}(1-q)^{n-1-m-\alpha} + (1-q) \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} q^{\alpha}(1-q)^{n-1-m-\alpha}\right]$$

$$EL = \frac{q^{m-\alpha_n}(1-q)^{\alpha_n}}{2}\left[(1-q) \sum_{\alpha \in Z_a^A} \binom{n-1-m}{\alpha} q^{n-1-m-\alpha}(1-q)^{\alpha} + q \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} q^{n-1-m-\alpha}(1-q)^{\alpha}\right]$$

The net increase in expected payoff of the final agent in $\hat{G}$ relative to G is given by:

$$\begin{aligned} EG - EL = &\frac{1}{2} \sum_{\alpha \in Z_a^A} \binom{n-1-m}{\alpha} \left[q^{\alpha_n+\alpha+1}(1-q)^{n-1-\alpha_n-\alpha} - q^{n-1-\alpha_n-\alpha}(1-q)^{\alpha_n+\alpha+1}\right] \\ &+ \frac{1}{2} \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} \left[q^{\alpha_n+\alpha}(1-q)^{n-\alpha_n-\alpha} - q^{n-\alpha_n-\alpha}(1-q)^{\alpha_n+\alpha}\right] \end{aligned}$$

If $Z_a^A = \varnothing$, the first sum equals zero; otherwise, it is non-negative. If $Z_b^A = \varnothing$, the second sum equals zero; otherwise it is strictly positive. Therefore, $EG - EL \geq 0$.

Case 2 Suppose $P(A \mid O_n) \in [1-q, q]$. Then, $x_n = A$ if $s_n = a$ and $x_n = B$ if $s_n = B$, and the final agent in G makes the correct decision only if they receive a correct signal. Therefore, a final agent in $\hat{G}$ experiences a gain if they make a correct choice despite having an incorrect signal. Conversely, they experience a loss if they receive a correct signal but make an incorrect choice. Thus,

$$\begin{aligned} EG = &\frac{q^{\alpha_n}(1-q)^{m-\alpha_n}}{2}\left[(1-q) \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} q^{\alpha}(1-q)^{n-1-m-\alpha}\right] \\ &+ \frac{q^{m-\alpha_n}(1-q)^{\alpha_n}}{2}\left[(1-q) \sum_{\alpha \in Z_a^B} \binom{n-1-m}{\alpha} q^{n-1-m-\alpha}(1-q)^{\alpha}\right] \\ EL = &\frac{q^{\alpha_n}(1-q)^{m-\alpha_n}}{2}\left[q \sum_{\alpha \in Z_a^B} \binom{n-1-m}{\alpha} q^{\alpha}(1-q)^{n-1-m-\alpha}\right] \\ &+ \frac{q^{m-\alpha_n}(1-q)^{\alpha_n}}{2}\left[q \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} q^{n-1-m-\alpha}(1-q)^{\alpha}\right] \end{aligned}$$

The net increase in expected payoff of the final agent in $\hat{G}$ relative to G is given by:

$$\begin{aligned} EG - EL = &\frac{1}{2} \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} \left[q^{\alpha_n+\alpha}(1-q)^{n-\alpha_n-\alpha} - q^{n-\alpha_n-\alpha}(1-q)^{\alpha_n+\alpha} \right] \\ &+ \frac{1}{2} \sum_{\alpha \in Z_a^B} \binom{n-1-m}{\alpha} \left[q^{n-1-\alpha_n-\alpha}(1-q)^{\alpha_n+\alpha+1} - q^{\alpha_n+\alpha+1}(1-q)^{n-1-\alpha_n-\alpha} \right] \end{aligned}$$

If $Z_b^A = \varnothing$, the first sum equals zero; otherwise it is strictly positive. If $Z_a^B = \varnothing$, the second sum equals zero; otherwise it is strictly positive. Therefore, $EG - EL \geq 0$.

Case 3 Suppose $P(A \mid O_n) \in (q, 1)$. For any $s_n \in \{a, b\}$, $x_n = A$ and the final agent in G makes the correct choice only if $\omega = A$. Therefore, the final agent experiences a gain in $\hat{G}$ if $\omega = B$ and $\hat{x}_n = B$. Conversely, they experience a loss if $\omega = A$ and $\hat{x}_n = B$. Thus,

$$EG = \frac{q^{m-\alpha_n}(1-q)^{\alpha_n}}{2} \left[(1-q) \sum_{\alpha \in Z_a^B} \binom{n-1-m}{\alpha} q^{n-1-m-\alpha}(1-q)^{\alpha} + q \sum_{\alpha \in Z_b^B} \binom{n-1-m}{\alpha} q^{n-1-m-\alpha}(1-q)^{\alpha} \right]$$

$$EL = \frac{q^{\alpha_n}(1-q)^{m-\alpha_n}}{2} \left[q \sum_{\alpha \in Z_a^B} \binom{n-1-m}{\alpha} q^{\alpha}(1-q)^{n-1-m-\alpha} + (1-q) \sum_{\alpha \in Z_b^A} \binom{n-1-m}{\alpha} q^{\alpha}(1-q)^{n-1-m-\alpha} \right]$$

The net increase in expected payoff of the final agent in $\hat{G}$ relative to G is given by:

$$\begin{aligned} EG - EL = &\frac{1}{2} \sum_{\alpha \in Z_a^B} \binom{n-1-m}{\alpha} \left[q^{n-1-\alpha_n-\alpha}(1-q)^{\alpha_n+\alpha+1} - q^{\alpha_n-\alpha+1}(1-q)^{n-1-\alpha_n-\alpha} \right] \\ &+ \frac{1}{2} \sum_{\alpha \in Z_b^B} \binom{n-1-m}{\alpha} \left[q^{n-\alpha_n-\alpha}(1-q)^{\alpha_n+\alpha} - q^{\alpha_n+\alpha}(1-q)^{n-\alpha_n-\alpha} \right] \end{aligned}$$

If $Z_a^B = \varnothing$ the first sum equals zero; otherwise it is strictly positive. If $Z_b^B = \varnothing$, the second sum equals zero; otherwise it is non-negative. Therefore, $EG - EL \geq 0$.

We conclude that, in every possible case, the final agent is at least as well off in network $\hat{G}$ as they are in G. Therefore, the expected payoff of the final agent is maximized by any network structure that is fully informative. □

Proof of Proposition 5

Consider agents i and j such that $|O_j| \leq 1$ and $|O_i| = 2$. The expected payoff of agent i is given by Equation (5) and the expected payoff of agent j is given by Equation (6). By the binomial theorem, these expected payoffs can equivalently be expressed as follows:

$$E_{S(n)}[u_j(x_j, \omega; G)] = qV$$
$$E_{S(n)}[u_i(x_i, \omega; G)] = q^2(3 - 2q)V$$

Then, it is straightforward to show the following:

$$E_{S(n)}[u_i(x_i, \omega; G)] - E_{S(n)}[u_j(x_j, \omega; G)] = q[q(3 - 2q) - 1]V$$

This difference is strictly positive since, by assumption, $q \in (1/2, 1)$. $\square$

Proof of Proposition 6

Fix n and q. Set $V = 1$, and let G be any arbitrary network structure. Consider any two agents i and j such that $i > j$ and $x_j \in O_i$ (i.e., the action of j, x_j, is observed by i). Note that agent i acts after j, observes every action j observes, and possibly observes additional actions. Let $O_{i-j} = O_i \setminus O_j$ denote the set of actions observed by agent i which are not observed by agent j. For any sequence of private signals, we define the following values, which describe the composition of information available to each agent.

$$\begin{aligned}
m_j &= |\{x_\ell \in O_j : P(A \mid O_\ell) \in [1-q, q]\}| \\
m_{ij} &= |\{x_\ell \in O_{i-j} : P(A \mid O_\ell) \in [1-q, q]\}| \\
\alpha_j &= |\{x_\ell \in O_j : x_\ell = A \text{ and } P(A \mid O_\ell) \in [1-q, q]\}| \\
\alpha_{ij} &= |\{x_\ell \in O_{i-j} : x_\ell = A \text{ and } P(A \mid O_\ell) \in [1-q, q]\}|
\end{aligned}$$

The social belief of agent i can then be expressed as follows

$$P(A \mid O_i) = \frac{q^{\alpha_j+\alpha_{ij}}(1-q)^{m_{ij}+m_j-\alpha_j-\alpha_{ij}}}{q^{\alpha_j+\alpha_{ij}}(1-q)^{m_{ij}+m_j-\alpha_j-\alpha_{ij}} + q^{m_{ij}+m_j-\alpha_j-\alpha_{ij}}(1-q)^{\alpha_j+\alpha_{ij}}}$$

Applying the decision rule derived in Proposition 1, the choice of agent i satisfies:

- If $s_i = a$,

$$x_i = \begin{cases} A, & \text{if } \alpha_{ij} \geq \frac{m_j+m_{ij}-1}{2} - \alpha_j \\ B, & \text{if } \alpha_{ij} < \frac{m_j+m_{ij}-1}{2} - \alpha_j \end{cases}$$

- If $s_i = b$,

$$x_i = \begin{cases} A, & \text{if } \alpha_{ij} > \frac{m_{ij}+m_j+1}{2} - \alpha_j \\ B, & \text{if } \alpha_{ij} \leq \frac{m_{ij}+m_j+1}{2} - \alpha_j \end{cases}$$

Define the following, where, by construction, $\alpha_{ij} \leq m_{ij}$ must also hold:

$$Z_a^A = \left\{ \alpha_{ij} : \alpha_{ij} \geq \frac{m_{ij} + m_j - 1}{2} - \alpha_j \right\} \qquad Z_b^A = \left\{ \alpha_{ij} : \alpha_{ij} > \frac{m_{ij} + m_j + 1}{2} - \alpha_j \right\}$$
$$Z_a^B = \left\{ \alpha_{ij} : \alpha_{ij} < \frac{m_{ij} + m_j - 1}{2} - \alpha_j \right\} \qquad Z_b^B = \left\{ \alpha_{ij} : \alpha_{ij} \leq \frac{m_{ij} + m_j + 1}{2} - \alpha_j \right\}$$

There are now multiple cases to consider. In each we will show that the expected gains (EG) are at least as large as the expected losses (EL) for agent i relative to agent j.

Case 1 Suppose $P(A \mid O_j) \in (0, 1-q)$. For any $s_j \in \{a, b\}$, $x_j = B$ and agent j makes the correct choice only if $\omega = B$. Therefore, agent i experiences a gain if $\omega = A$ and $x_i = A$. Conversely, they experience a loss if $\omega = B$ and $x_i = A$. Thus,

$$EG = \frac{q^{\alpha_j}(1-q)^{m_j - \alpha_j}}{2} \left[q \sum_{\alpha \in Z_a^A} \binom{m_{ij}}{\alpha} q^{\alpha}(1-q)^{m_{ij}-\alpha} + (1-q) \sum_{\alpha \in Z_b^A} \binom{m_{ij}}{\alpha} q^{\alpha}(1-q)^{m_{ij}-\alpha} \right]$$
$$EL = \frac{q^{m_j - \alpha_j}(1-q)^{\alpha_j}}{2} \left[(1-q) \sum_{\alpha \in Z_a^A} \binom{m_{ij}}{\alpha} q^{m_{ij}-\alpha}(1-q)^{\alpha} + q \sum_{\alpha \in Z_b^A} \binom{m_{ij}}{\alpha} q^{m_{ij}-\alpha}(1-q)^{\alpha} \right]$$

Then, the net increase in expected payoff of agent i relative to agent j is given by:

$$\begin{aligned} EG - EL = & \frac{1}{2} \sum_{\alpha \in Z_a^A} \binom{m_{ij}}{\alpha} \left[a^{\alpha + \alpha_{ij} + 1}(1-q)^{m_{ij} + m_j - \alpha - \alpha_j} - q^{m_{ij} + m_j - \alpha - \alpha_{ij}}(1-q)^{\alpha + \alpha_j + 1} \right] \\ & + \frac{1}{2} \sum_{\alpha \in Z_b^A} \binom{m_{ij}}{\alpha} \left[q^{\alpha + \alpha_j}(1-q)^{m_{ij} + m_j + 1 - \alpha - \alpha_j} - q^{m_{ij} + m_j + 1 - \alpha - \alpha_j}(1-q)^{\alpha + \alpha_j} \right] \end{aligned}$$

If $Z_a^A = \varnothing$, the first sum equals zero; otherwise, it is non-negative. If $Z_b^A = \varnothing$, the second sum equals zero; otherwise it is strictly positive. Therefore, $EG - EL \geq 0$.

Case 2 Suppose $P(A \mid O_j) \in [1-q, q]$. Then, $x_j = A$ if $s_j = a$ and $x_j = B$ if $s_j = B$, and agent j makes the correct choice only if they receive a correct signal. Therefore, agent i experiences a gain if they make a correct choice despite having an incorrect signal. Conversely, they experience a loss if they receive a correct signal but make an incorrect choice. Thus,

$$EG = \frac{q^{\alpha_j}(1-q)^{m_j-\alpha_j}}{2}\left[(1-q)\sum_{\alpha\in Z_b^A}\binom{m_{ij}}{\alpha}q^{\alpha}(1-q)^{m_{ij}-\alpha}\right]$$
$$+\frac{q^{m_j-\alpha_j}(1-q)^{\alpha_j}}{2}\left[(1-q)\sum_{\alpha\in Z_a^B}\binom{m_{ij}}{\alpha}q^{m_{ij}-\alpha}(1-q)^{\alpha}\right]$$
$$EL = \frac{q^{\alpha_j}(1-q)^{m_j-\alpha_j}}{2}\left[q\sum_{\alpha\in Z_a^B}\binom{m_{ij}}{\alpha}q^{\alpha}(1-q)^{m_{ij}-\alpha}\right]$$
$$+\frac{q^{m_j-\alpha_j}(1-q)^{\alpha_j}}{2}\left[q\sum_{\alpha\in Z_b^A}\binom{m_{ij}}{\alpha}q^{m_{ij}-\alpha}(1-q)^{\alpha}\right]$$

Then, the net increase in expected payoff of agent i relative to agent j is given by:

$$EG-EL = \frac{1}{2}\sum_{\alpha\in Z_b^A}\binom{m_{ij}}{\alpha}\left[q^{\alpha+\alpha_j}(1-q)^{m_{ij}+m_j-\alpha-\alpha_j} - q^{m_{ij}+m_j+1-\alpha-\alpha_j}(1-q)^{\alpha+\alpha_j}\right]$$
$$+\frac{1}{2}\sum_{\alpha\in Z_a^B}\binom{m_{ij}}{\alpha}\left[q^{m_{ij}+m_j-\alpha-\alpha_j}(1-q)^{\alpha+\alpha_j+1} - q^{\alpha_j+\alpha+1}(1-q)^{m_{ij}+m_j-\alpha-\alpha+j}\right]$$

If $Z_b^A=\varnothing$, the first sum equals zero; otherwise it is strictly positive. If $Z_a^B=\varnothing$, the second sum equals zero; otherwise it is strictly positive. Therefore, $EG-EL\geq 0$.

Case 3 Suppose $P(A\mid O_j)\in(q,1)$. For any $s_j\in\{a,b\}$, $x_j=A$ and agent j makes the correct choice only if $\omega=B$. Therefore, agent i experiences a gain if $\omega=B$ and $x_i=B$. Conversely, they experience a loss if $\omega=A$ and $x_i=B$. Thus,

$$EG = \frac{q^{m_j-\alpha_j}(1-q)^{\alpha_j}}{2}\left[(1-q)\sum_{\alpha\in Z_a^B}\binom{m_{ij}}{\alpha}q^{m_{ij}-\alpha}(1-q)^{\alpha} + q\sum_{\alpha\in Z_b^B}\binom{m_{ij}}{\alpha}q^{m_{ij}-\alpha}(1-q)^{\alpha}\right]$$
$$EL = \frac{q^{\alpha_j}(1-q)^{m_j-\alpha_j}}{2}\left[q\sum_{\alpha\in Z_a^B}\binom{m_{ij}}{\alpha}q^{\alpha}(1-q)^{m_{ij}-\alpha} + (1-q)\sum_{\alpha\in Z_b^B}\binom{m_{ij}}{\alpha}q^{\alpha}(1-q)^{m_{ij}-\alpha}\right]$$

Then, the net increase in expected payoff of agent i relative to agent j is given by:

$$EG-EL = \frac{1}{2}\sum_{\alpha\in Z_a^B}\binom{m_{ij}}{\alpha}\left[q^{m_{ij}+m_j-\alpha-\alpha_j}(1-q)^{\alpha+\alpha_j+1} - q^{\alpha+\alpha_j+1}(1-q)^{m_{ij}+m_j-\alpha-\alpha_j}\right]$$
$$+\frac{1}{2}\sum_{\alpha\in Z_b^B}\binom{m_{ij}}{\alpha}\left[q^{m_{ij}+m_j+1-\alpha-\alpha_j}(1-q)^{\alpha+\alpha_j} - q^{\alpha+\alpha_j}(1-q)^{m_{ij}+m_j+1-\alpha-\alpha_j}\right]$$

If $Z_a^B=\varnothing$, the first sum equals zero; otherwise, it is strictly positive. If $Z_b^B=\varnothing$, the second sum equals zero; otherwise, it is non-negative. Therefore, $EG-EL\geq 0$.

We conclude that, in every possible case, the expected payoff of agent i is at least as large as that of agent j. Therefore, every agent is weakly better off in expectation than any agent whose action they observe. $\square$

Proof of Proposition 7

Fix n and q, set $V = 1$, let G^C denote the complete history network, and let $\hat{G}$ denote a fully informative network. We can express aggregate expected payoffs in $\hat{G}$ as follows:

If n is odd,

$$E[U;\hat{G}] = (n-1)q + q\binom{n-1}{\frac{n-1}{2}}\left[q(1-q)\right]^{\frac{n-1}{2}} + \sum_{r=\frac{n+1}{2}}^{n-1}\binom{n-1}{r}q^r(1-q)^{n-1-r}$$

If n is even,

$$E[U;\hat{G}] = (n-1)q + q\binom{n-1}{\frac{n}{2}-1}q^{\frac{n}{2}-1}(1-q)^{\frac{n}{2}} + q\binom{n-1}{\frac{n}{2}}q^{\frac{n}{2}}(1-q)^{\frac{n}{2}-1} + \sum_{r=\frac{n}{2}+1}^{n-1}\binom{n-1}{r}q^r(1-q)^{n-1-r}$$

In G^C, we can express the aggregate expected payoffs in G^C as follows (where the index r counts pairs of subsequent agents which alternate actions; i.e., this describes how long cascade formation is delayed).

If n is odd,

$$E[U;G^C] = \sum_{r=0}^{\frac{n-3}{2}}\left[2q(1-q)\right]^r\left[q^2(n-r) + (1-q)^2 r\right] + \left[2q(1-q)\right]^{\frac{n-1}{2}}\left(\frac{n-1}{2} + q\right)$$

If n is even,

$$E[U;G^C] = \sum_{r=0}^{\frac{n}{2}-1}\left[2q(1-q)\right]^r\left[q^2(n-r) + (1-q)^2 r\right] + \left[2q(1-q)\right]^{\frac{n}{2}}\left(\frac{n}{2}\right)$$

Subsequent analysis is simplified by establishing a lower bound on the aggregate expected payoffs in G^C. Proposition 6 implies that all agents after the third must have expected payoff at least as large as the third agent who has expected payoff equal to $q^2 + 2q^2(1-q)$. Closer inspection reveals that, beginning with the fifth agent they are actually strictly better in expectation than the third agent. So, assuming $n \geq 5$, we can establish the following bound on aggregate expected payoffs in G^C:

$$E[U;G^C] > 2q + (n-2)[q^2 + 2q^2(1-q)] = 2q + (n-2)q^2(3-2q)$$

In the case of $n = 3$, Proposition 3, implies that $E[U;G^C] - E[U;\hat{G}] = 0$, since the payoff of non-final agents are necessarily q in both cases. For the case of $n = 4$, it is straightforward to verify that $E[U;G^C] - E[U;\hat{G}] > 0$. Therefore, we will move on to the general case of $n \geq 5$. We will examine the cases of n odd and n even separately; however, regardless of whether n is odd or even, the (bounded) difference in aggregate expected payoffs is given by $\varphi - E[u_n;\hat{G}]$, where $\varphi = (n-2)q^2(3-2q) - (n-3)q$. Finally note that, by the binomial theorem, we can rewrite φ as follows:

$$\varphi = \varphi \sum_{r=0}^{n-1} \binom{n-1}{r} q^r (1-q)^{n-1-r}$$

Case 1 Suppose n is odd. Then, we have the following:

$$E[U;G^C] - E[U;\hat{G}] > \varphi - q\binom{n-1}{\frac{n-1}{2}}[q(1-q)]^{\frac{n-1}{2}} - \sum_{r=\frac{n+1}{2}}^{n-1} \binom{n-1}{r} q^r (1-q)^{n-1-r}$$

Now, for any $\lambda \in [0,1]$, note that:

$$\sum_{r=\frac{n+1}{2}}^{n-1} \binom{n-1}{r} q^r (1-q)^{n-1-r} = \lambda \sum_{r=0}^{\frac{n-3}{2}} \binom{n-1}{r} q^{n-1-r}(1-q)^r + (1-\lambda) \sum_{r=\frac{n+1}{2}}^{n-1} \binom{n-1}{r} q^r (1-q)^{n-1-r}$$

Now, we can express the (bounded) difference in aggregate payoffs as follows:

$$\begin{aligned}
E[U;G^C] - E[U;\hat{G}] > &\sum_{r=0}^{\frac{n-3}{2}} \binom{n-1}{r} \left[q^r(1-q)^{n-1-r}\varphi - q^{n-1-r}(1-q)^r\lambda\right] \\
&+ \binom{n-1}{\frac{n-1}{2}}[q(1-q)]^{\frac{n-1}{2}}\left[\varphi - q\right] \\
&+ \sum_{r=\frac{n+1}{2}}^{n-1} \binom{n-1}{r} q^r(1-q)^{n-1-r}\left[\varphi - (1-\lambda)\right]
\end{aligned}$$

We first consider the second term which is non-negative if and only if $\varphi - q \geq 0$.

$$\begin{aligned}\varphi - q &= (n-2)q^2(3-2q) - (n-2)q \\ &= (n-2)q\left[q(3-2q) - 1\right]\end{aligned}$$

This is strictly positive for $q \in (1/2, 1)$. Terms in the second summation are non-negative if and only if $\lambda \geq 1 - \varphi$ and terms in the first summation are non-negative if and only if

$$\lambda \leq \varphi \left[\frac{1-q}{q}\right]^{n-1-2r}$$

Therefore, we must choose a weight, λ, to satisfy both constraints. The proof of this case is completed by demonstrating that such a weight exists. We select λ to satisfy the tightest constraint on the first summation, which occurs for $r = \frac{n-3}{2}$. That is, $\lambda = \varphi[\frac{1-q}{q}]^2$. Then, our second condition requires that $\varphi\left[\frac{1-q}{q}\right]^2 \geq 1 - \varphi$. This is equivalent to $\varphi(1-q)^2 \geq (1-\varphi)q^2$, which holds since $q \in (1/2, 1)$ and $\varphi > q$. Therefore, it is possible to choose a weight, λ, to satisfy the preceding analysis. We conclude that for $n \geq 5$ and odd aggregate expected payoffs are strictly higher in G^C than in $\hat{G}$.

Case 2 Suppose n is even. Then, we have the following:

$$\begin{aligned}E[U; G^C] - E[U; \hat{G}] > \varphi - q\binom{n-1}{\frac{n}{2}-1}q^{\frac{n}{2}-1}(1-q)^{\frac{n}{2}} + q\binom{n-1}{\frac{n}{2}}q^{\frac{n}{2}}(1-q)^{\frac{n}{2}-1} \\ + \sum_{\frac{n}{2}+1}^{n-1}\binom{n-1}{r}q^r(1-q)^{n-1-r}\end{aligned}$$

Now, for any $\lambda \in [0, 1]$, note that:

$$\sum_{\frac{n}{2}+1}^{n-1}\binom{n-1}{r}q^r(1-q)^{n-1-r} = \lambda\sum_{r=0}^{\frac{n}{2}-2}\binom{n-1}{r}q^{n-1-r}(1-q)^r + (1-\lambda)\sum_{\frac{n}{2}+1}^{n-1}\binom{n-1}{r}q^r(1-q)^{n-1-r}$$

Then, we can express the (bounded) difference in aggregate payoffs as follows:

$$\begin{aligned}
E[U;G^C]-E[U;\hat{G}] &= \sum_{r=0}^{\frac{n}{2}-2}\binom{n-1}{r}\left[q^r(1-q)^{n-1-r}\varphi-q^{n-1-r}(1-q)^r\lambda\right] \\
&+\binom{n-1}{\frac{n}{2}-1}q^{\frac{n}{2}-1}(1-q)^{\frac{n}{2}}\left[\varphi-q\right]+\binom{n-1}{\frac{n}{2}}q^{\frac{n}{2}}(1-q)^{\frac{n}{2}-1}\left[\varphi-q\right] \\
&+\sum_{\frac{n}{2}+1}^{n-1}\binom{n-1}{r}q^r(1-q)^{n-1-r}\left[\varphi-(1-\lambda)\right]
\end{aligned}$$

Recall that, for $q \in (1/2, 1)$, $\varphi > q$, and so, the second and third terms are strictly positive. Terms in the second summation are non-negative if and only if $\lambda \geq 1-\varphi$ and terms in the first summation are non-negative if and only if

$$\lambda \leq \varphi\left[\frac{1-q}{q}\right]^{n-1-2r}$$

Therefore, we must choose a weight, λ, to satisfy both constraints. The proof of this case is completed by demonstrating that such a weight exists. We select λ to satisfy the tightest constraint on the first summation, which occurs for $r=\frac{n}{2}-2$. That is, $\lambda=\varphi\left[\frac{1-q}{q}\right]$. Then, our second condition requires that $\varphi\left[\frac{1-q}{q}\right] \geq 1-\varphi$. This is equivalent to $\varphi(1-q) \geq (1-\varphi)q$, which holds since $q \in (1/2, 1)$ and $\varphi > q$. Therefore, it is possible to choose a weight, λ, to satisfy the preceding analysis. We conclude that for $n \geq 6$ and even aggregate expected payoffs are strictly higher in G^C than in $\hat{G}$. $\square$

C Online Appendix: Experiment Documentation

C.1 Experimental Design

Table 13: Signal Sequence Frequencies (by Group)

Sequence	Frequency	Percent	Sequence	Frequency	Percent
AAAAA	21	8.57	ABAAA	2	0.82
AAAAB	8	3.27	ABAAB	5	2.04
AAABA	12	4.90	ABABA	15	6.12
AAABB	10	4.08	ABABB	5	2.04
AABAA	10	4.08	ABBAA	10	4.08
AABAB	5	2.04	ABBAB	3	1.22
AABBA	9	3.67	ABBBA	15	6.12
BBAAB	7	2.86	ABBBB	20	8.16
BBABA	13	5.31	BAABA	15	6.12
BBABB	10	4.08	BABAB	10	4.08
BBBAA	3	1.22	BABBA	5	2.04
BBBBA	10	4.08	BBAAA	5	2.04
BBBBB	17	6.94	-	-	-

C.2 Sample Experimental Instructions (N3)

Welcome to the experiment. Please put away your phones, books, etc. for the duration of the experiment. We will need your full attention. From this point on, please also do not communicate with other subjects in the experiment.

These instructions will describe the experiment, your decisions, and payoffs. Your earnings for this experiment may depend on your decisions, decisions of others, and chance. You will receive $7 for arriving on time and participating in the experiment, and will have the opportunity to earn additional money throughout the course of the experiment. The $7 show-up payment plus any money that you earn during the experiment will be paid to you, in cash, at the end of the experiment.

If you have a question at any point please raise your hand and your question will be answered.

Groups The first part of the experiment will consist of 7 decision making rounds. At the beginning of every round, all participants will be randomly matched into groups of 5. You will not know who is in your group in any round, just that these are 4 other participants in the experiment. Groups will be randomly re-matched after every round.

Event Every round, one of the two events, A or B, will take place. Each event is equally likely to happen; in other words there is a 50-50 chance for either A or B to occur. This random process (you can think of it as a coin toss) will determine which event takes place at the beginning of every round.

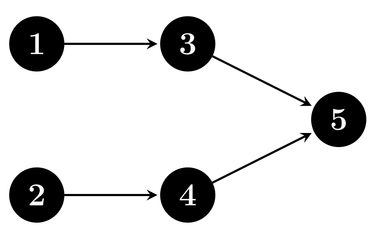
Signal Every person in the group will independently get their signal of the event, A or B. The signal precision is 0.7 or 70%. This means that there is 70% chance the signal you receive corresponds to the true event A or B, and 30% chance your signal does not correspond to the true event. Every person in the group draws their own signal. You will see your signal but not the signals of other persons in your group. Again, signal precision of 0.7 or 70% means that on average, out of many draws, about 70% of signals are expected to correspond to the true event (A or B), and 30% of signals are expected to indicate the other event. Since the signals are drawn randomly various compositions of true and false signals are possible but the likelihood of your signal being the same as the true event is 0.7 or 70%.

Are there any questions?

Choice Every person in the group has to make a choice, A or B, for what is the true event. The true event, A or B, will be announced at the end of the round.

Payoff Every person whose choice corresponds to the true event, gets a payoff of $9.

Sequence All people in the group will make their choices, A or B, in a sequence. The structure of the sequence is shown below.



There are 5 positions within a sequence, and each person in a group will be randomly assigned to one of the positions at the beginning of every round. Your position within the sequence also indicates what information you have available to you when you make your choice of the true event, A or B. Specifically, in addition to your private signal, you observe choices of persons in your group that were in earlier positions as indicated by the arrows. The order in which all group participants make decisions are as the number of the position. The individual in position 1 will make their decision first, and will see only their signal. Then person in position 2 makes their choice and observes only their signal. Person in position 3 will observe their private signal and the choice of individual in position 1 before making a decision. Person in position 4 observes own signal and the choice of individual in position 2, and makes own choice about the state. Person in position 5 will observe their signal, and the choices of the 4 individuals in the group who made decisions in earlier positions before making own choice.

Earnings Your earnings in a round are $9.00 if your choice matches the true event, A or B, and $0 otherwise. After the completion of 7 decision rounds, one round payoff will be randomly drawn for payment to be added to your $7 show up fee. Each round is equally likely to be drawn for payment.

To Summarize At the beginning of every round, all subjects will be randomly re-matched into new groups of 5 persons each. Every person in a group is randomly assigned a position within a sequence structure. Every person draws a private signal of the true event, A or B, with signal precision, or accuracy, being 0.7 or 70%. At the time of making a decision, A or B, you may also observe decisions of persons in positions prior to yours as indicated by the sequence structure. After all persons within a sequence have made their choice, the true event will be revealed. If your choice matches the true event, your payoff in a round is $9. If your choice does not match the true event, your payoff in a round is $0. Since groups and positions are assigned randomly at the beginning of every round, the compositions of which persons are in the group with you, and your positions within a sequence, are likely to change from round to round. Your earnings from this part of the experiment will be your payoff from one randomly drawn round.

Are there any questions?

C.3 Experiment Screenshots

Figure 6: Individual Decision Screen (N3, Position 5)

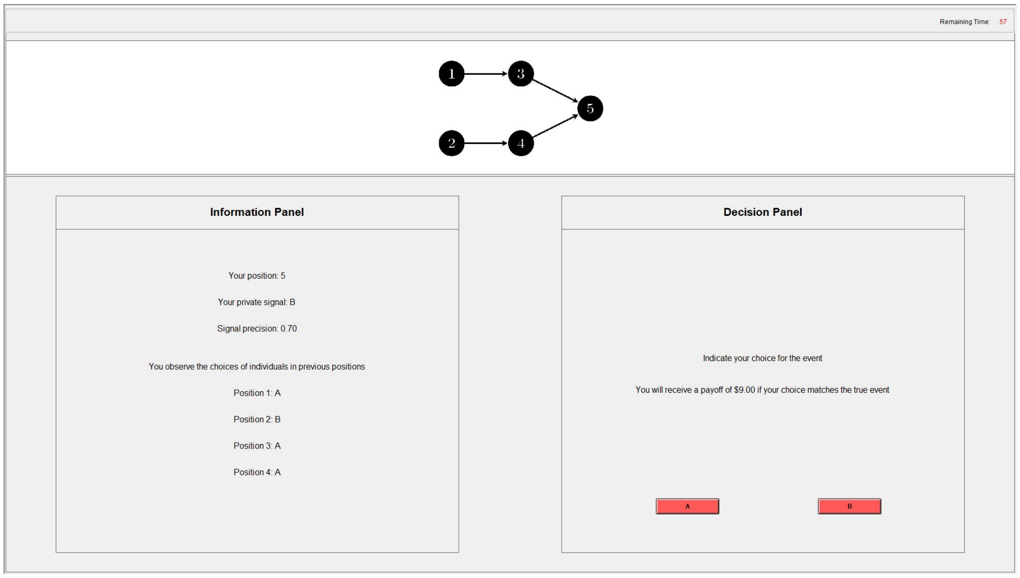


Figure 7: Endogenous Position Auction Screen